Molecular techniques for rhizobium identification

Dooley, John J.

Award date:
1997

Awarding institution:
University of Bath

Molecular Techniques for *Rhizobium* Identification.

submitted by John J. Dooley

for the Degree of PhD

of the University of Bath

1996.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Dedicated to:

My friends -

for being just that.

Acknowledgements

I would like to thank the following people for their kind help and assistance during the years I have spent researching and writing this thesis. Those not named have been overlooked through my forgetfulness but remain thanked none the less.

I suppose I should, as all good PhD students should, firstly thank my supervisor Dr. S. P. Harrison. I thank him for leaving me pretty much to my own devices in the lab which has enabled me to develop my skills (sometimes through trial and error) without someone breathing down my neck. However, he has always been there (or there abouts) to show me how things should (or shouldn't when it comes to general DIY repairs) be done.

I would also like to thank Dr. J.R. Beeching, who as my Bath University supervisor, kept my work up to standard and provided guidance during my writing up stage. I would also like to thank him for his assistance with statistical problems.

Thanks are also sent to my colleagues, Dr. L. Mytton, Dr. M. Dye, Dr. A. Cresswell and Dr. L. Skøt, at the Welsh Plant Breeding Station, Aberyswyth, who have been involved with this project. Thanks are also offered to the WPBS which partially sponsored this research and which provided the majority of the isolates employed.

I should next like to thank 'My Mate' Nicky (Nicola Cosser) who, managed to maintain my sanity during my period of writing up. Thanks must also go to the other postgrads who work at the Royal Agricultural College; To Vanja Karamatic, my fellow geneticist (for managing to last nearly as long as I have in the genetics lab), to Andrew Thompson (for being a fellow postgrad in the early years) and to Kate Warman (for Saturday lunch-time sessions in the 'Tunnel House' when we were both working over the weekend). Thanks are also offered to my fellow postgrads and friends who have made my stay here enjoyable, Graham Smith, Leanne Parker, Adrian Crew, Sybil Ellis, Alexia Buckwell and Mark Lovatt, Little Jo and Dicky (Jo and Richard Blunden).

Thanks also go to the following who have assisted me during the last five or so years:

The lab technicians Sylvia Duncan, John Stevens and Jean Taylor for their help with technical problems, orders and equipment.

The library staff especially Jane Grogan, Helen Thompson and recently Evelyn Santer, who handled my ever constant stream of inter-library loan requests.

Mr Julian Swindell for his help with the graphical presentation used in this report.

Dr Mike Gooding for his assistance with GENSTAT programming.

The ground and kitchen staff for their friendship and gossip over the past years. Special thanks also go to Paul Collett for Saxophone lessons over the last two years which has kept me from going totally insane.

The "squash club" - Gareth Thomas, Steve Finch, John Nixon, Nigel Bridges, Mary Cox, Chris Stevens, John Alliston and all the other players who rounded off each week with some exercise and fun.

I should like to thank all those members of The Royal Agricultural College who have made my stay here enjoyable by their general friendliness and help.

Last, but not least, I would like to thank my parents, brothers and sister for their support over the last years and for their nagging me “to get written up".

Contents

Abbreviations

15mer	-	15 bp DNA Primer.
16S rRNA	-	16S ribosomal RNA genes.
23S rRNA	-	23S ribosomal RNA genes.
Ab	-	Antibody.
AMOVA	-	Analysis of Molecular Variance.
ARDRA	-	Amplified Ribosomal DNA Restriction Analysis.
bp	-	Base-pair.
Br	-	*Bradyrhizobium* isolate.
bv	-	Biovar.
Dig	-	Digoxigenin Moiety.
Dig-dUTP	-	Digoxigenin Labelled dUTP.
dATP	-	Deoxyadenosine-Triphosphate.
dCTP	-	Deoxycytidine-Triphosphate.
dGTP	-	Deoxyguanosine-Triphosphate.
dNTP	-	Deoxyribonucleoside-Triphosphate.
dTTP	-	Deoxythymidine-Triphosphate.
dUTP	-	Deoxyuridine-Triphosphate.
DNA	-	Deoxyribonucleic Acid.
EDTA	-	Ethylenediamine Tetra-acetic Acid - Disodium Salt.
ELISA	-	Enzyme Linked Immunosorbent Assay.
ET	-	Electrophoretic Type.
HOMOVA	-	Homogeneity of Molecular Variance.
IAR	-	Intrinsic Antibiotic Resistance.
IS	-	Insertion Sequence.
Kb	-	Kilobase-pairs.
µl	-	Micro-Litre.

Lumigen-PPD	-	3-(2'-Spiroadamantane)-4-methoxy-4-(3"-phosphoryloxy)-phenyl-1,2-dioxetane.
µM	-	Micro-Molar.
M	-	Molar.
M1A	-	*R. meliloti* Species-specific Probe.
M1B	-	*R. meliloti* Species-specific Probe.
M1C	-	*R. meliloti* Species-specific Probe.
mg	-	Milli-Gram.
ml	-	Milli-Litre.
ml^{-1}	-	per Milli-Litre.
MLO	-	Mycoplasma Like Organism.
mM	-	Milli-Molar.
MPN	-	Most Probable Number.
nif	-	Nitrogen Fixation Genes.
nod	-	Nodulation Genes.
PA	-	Peptone Based Growth Media.
PAGE	-	Polyacrylamide Gel Electrophoresis.
PAUP	-	Phylogenetic Analysis Using Parsimony.
PCA	-	Principal Components Analysis.
PCO	-	Principal Coordinates Analysis.
PCR	-	Polymerase Chain Reaction.
RAPD	-	Randomly Amplified Polymorphic DNA.
RCR	-	Rothamsted Collection of *Rhizobium*.
RFLP	-	Restriction Fragment Length Polymorphism.
Rm	-	*Rhizobium meliloti* isolate.
Rp	-	*R. leguminosarum* bv *phaseoli* isolate.
RS	-	Repeat Sequence.
Rt	-	*R. leguminosarum* bv *trifolii* isolate.
Rv	-	*R. leguminosarum* bv *viciae* isolate.

SDS	-	Sodium Dodecyl Sulphate.
SDW	-	Sterile Distilled Water.
SPH1	-	Random DNA Primer - (GAC_5).
SPH3	-	Random DNA Primer - (GACGACAGCGGC).
SPH7	-	Random DNA Primer - (CAGCCACAGCGC).
SSC	-	Saline Sodium Citrate.
Sym	-	Symbiotic Plasmid.
T1E	-	*R. leguminosarum* bv *trifolii* Species-specific Probe.
T37-3	-	*R. leguminosarum* bv *trifolii* Strain-specific Probe.
TE	-	Tris-EDTA Buffer.
TG	-	Transgoed soil isolate - not identified.
Tn5	-	Transposon 5.
Tris-Base	-	Tris(hydroxymethyl)aminomethane.
Tris-Cl	-	Tris(hydroxymethyl)aminomethane Hydrochloride.
TY	-	Tryptone and Yeast Based Growth Media.
UPGMA	-	Unweighted Pair Group Method with Analysis.
V1D	-	*R. leguminosarum* bv *viciae* Species-specific Probe.
v/v	-	Volume per Volume.
w/v	-	Weight per Volume.

Abstract

The production of RAPD fingerprints from a selection of 84 *Rhizobium* and *Bradyrhizobium* isolates was possible using three random primers, SPH1, SPH3 and SPH7, in single and double primed amplification reactions. Statistical analysis (PCO and Cluster Analysis) of the RAPDs permitted classification of the strains into their respective species. Primer SPH1 produced the best fingerprints for strain classification although the use of SPH3 and SPH7, in a double primed reaction, resulted in a better separation of *R. leguminosarum* biovars *trifolii* and *viciae*.

A series of potential species- and strain-specific bands were identified from the RAPD fingerprints and labelled with the non-isotopic Digoxigenin marker. These probes were used to study rhizobial isolates by hybridisation to colony blots and Southern blots of restriction digested genomic DNA. Results generally concurred with those from RAPD analysis but also indicated that none of the probes were in fact species-specific. Only one strain-specific probe was identified as useful for these purposes although results indicated that several others may be of use for species-specific studies.

It was concluded that RAPD fingerprinting is of use for taxonomic studies of *Rhizobium* and that it can provide a method of producing species- and strain-specific probes for further studies of this genera.

General Introduction

General Introduction

General Background:

In the area of plant-microbe interactions, the symbiotic relationship between the leguminous plants and the bacteria of the genera *Rhizobium* and *Bradyrhizobium* has a special significance. The leguminous plants include alfalfa, beans, clover and peas, all of which are of economic and agricultural importance. When a leguminous plant is infected with an appropriate *Rhizobium* or *Bradyrhizobium* the resultant root nodule has the ability to fix gaseous nitrogen. Levels of nitrogen fixation have been calculated at between 50 and 300 Kg of nitrogen fixed hectare^{-1} year^{-1} depending on the species of *Rhizobium* employed. This is over 100 times more nitrogen fixed hectare^{-1} year^{-1} than is obtained from free living species such as those of the family Azobacteraceae (Campbell, 1985). This value of the legume-*Rhizobium* symbiotic relationship is of great agricultural significance, especially in areas with poor, unfertilised soils.

Rhizobium Microbiology:

Species of the genera *Rhizobium* and *Bradyrhizobium* are soil-dwelling bacteria which are capable of nodulating leguminous plants and fixing nitrogen. The bacteria exist in two states, the free-living state and the bacteroid state. The free-living bacteria are found in soil whilst the larger (about 2 μm) bacteroid state is found only in nodules. Only the bacteroid is capable of fixing nitrogen.

The formation of root nodules on leguminous plants is controlled by chemical interactions between both the plant and the *Rhizobium* bacteria. The host plant and *Rhizobium* strain are attracted to each other by these chemical interactions, and by the use of non-polar flagella the *Rhizobium* isolates move toward the plant root hairs. The *Rhizobium* initiates the formation of a nodule by invading the root hair. Invasion of the root hair causes a localised cell growth in that region which results in the formation of a nodule. The invading bacteria then multiply to fill the available space before nitrogen fixation starts. Active nitrogen fixing nodules can be identified by their pinkish hue

which is caused by the presence of leghaemoglobin, an oxygen carrier which protects the oxygen-sensitive, nitrogen-fixing, nitrogenase enzyme.

Both nodulation and nitrogen fixation are under the control of genes (*nod* and *nif* respectively) which are located on the Sym (Symbiotic) plasmid. A size variation between Sym plasmids from various *Rhizobium* species has been reported. DeJong *et al.*, (1982), who transferred Sym plasmids between isolates in order to improve symbiotic properties, report finding Sym plasmids of 190 MDaltons (pRL6JI) and 220 MDa (pRL10JI) in *R. leguminosarum* (now *R. leguminosarum* bv *viciae*) isolates. Soberón-Chávez *et al.*, (1986), who studied a Sym plasmid from *R. leguminosarum* bv *phaseoli* report its size to be about 400 Kb.

The Sym plasmid is only one of a collection of large plasmids carried by *Rhizobium* species. These range in size between 50 MDa and 1000MDa, with the numbers per strain varying between 1 and about 8 (Djordevic *et al.*, 1982; Prakash and Atherley, 1986). Some of these plasmids remain cryptic, i.e. their function is unknown. The large amounts of plasmid DNA observed in some strains can amount to about 25% of the total genomic DNA content, this having been estimated at about 5.4-7.6 $x10^9$ Da (Chakraborti *et al.*, 1983). Estimates of 5379 Kb (*R. meliloti*) and 6195 Kb (*B. japonicum*) have also been made by Sobral *et al.*, (1991) who calculated these sizes from DNA fragment sizes resulting from DNA digestion with *Spe*I, a rare cutting enzyme.

The formation of the symbiotic legume-*Rhizobium* relationship is not achieved randomly, but by a highly selective method. Only a small, specific number of *Rhizobium* or *Bradyrhizobium* are able to form effective nodules with each legume type. The classification of *Rhizobium* was, therefore, originally based on the ability of the bacteria to form nodules with specific plants. Recently it has been found that certain legumes such as beans (*Phaseolus*) are somewhat promiscuous in their nodule formation with *Rhizobium* isolates. This has resulted in an incorrect classification of some strains. In order to correctly classify the *Rhizobium* and *Bradyrhizobium* it is necessary to understand the basic concepts of bacterial classification.

Bacterial Classification:

Bacterial classification involves the grouping of micro-organisms based on their relationship to other micro-organisms. This process follows strict taxonomic rules, but, once established it should be possible to identify new organisms and to categorise them within the classification system.

In order to classify any biological system it is first necessary to define a basic taxonomic unit. The accepted unit for classification purposes is the species. When studying higher organisms species are readily identifiable by their reproductive isolation, however, bacterial species are not so easily defined. A good working definition of a bacterial species is "a group of strains that show a high degree of overall phenotypic similarity and that differ from related strain groups with respect to many independent characteristics" (Stanier *et al.*, 1987).

Genetic variation within micro-organisms permits sub-divisions, known as subspecies, of the species to be defined. It is often desirable to identify or define subspecies or individual strains from a particular species. For example certain strains from a species produce toxins or are pathogenically virulent whereas other members of this species are harmless. In this case it would be useful to be able to identify the harmful strains so counter-measures to their pathogenicity or toxicity can be developed. In the case of *Rhizobium* it is useful to be able to identify those strains which are efficient nitrogen fixers (*nod*$^{+}$) or conversely those that are incapable of nitrogen fixing (*nod*$^{-}$). Obviously those strains which lack the ability to fix nitrogen would be of no use as inoculum whereas those which are effective fixers would be more suited to this purpose.

Traditional methods of bacterial classification include Gram reactions, morphology, ability to utilise various carbon, nitrogen and sulphur sources, symbiotic relationships and antibiotic sensitivity. Using these and other methods *Rhizobium* and *Bradyrhizobium* have been identified as belonging to the taxonomic group of Gram-negative, aerobic rods and cocci. This large taxonomic group consists of several other

bacterial families including the Pseudomonadaceae, Legionellaceae, and the Azobacteraceaea. Many species of the first two families are pathogenic and include *Pseudomonas aeruginosa* (a human pathogen), the *Xanthomonas* species (all of which are plant pathogens) and *Legionella pneumophila* (the causal agent of Legionnaires diseases). This pathogenic ability is observed in the genera of the Rhizobiaceae family such as species of *Agrobacterium*, which are plant pathogens, and the species of *Rhizobium* and *Bradyrhizobium* which can be regarded as specialised plant pathogens. *Agrobacterium* species, however, are unable to fix nitrogen whereas species of *Rhizobium* and *Bradyrhizobium* are capable of this when in association with an appropriate leguminous plant. Species of the family Azobacteraceae are also capable of nitrogen fixation although this is achieved when the bacteria are in a free living state. These bacteria, in fact, have no pathogenic tendencies and always remain as free living soil bacteria.

Although *Rhizobium*, *Bradyrhizobium* and the Azobacteraceae account for a large number of the nitrogen fixing bacteria they are not the only ones capable of this. The other major group of nitrogen fixers are actinomycetes of the genus *Frankia*. These bacteria, like the *Rhizobium*, fix nitrogen when in symbiosis with plants, however, these are non-leguminous plants and include trees such as Alder (*Alnus*) and also pioneering plants which colonise poor soils. Little work has been carried out on these bacteria although it is believed that they infect the roots of their respective plants in a manner similar to that seen in *Rhizobium*-legume interactions. However, it is known that the *Frankia*-non-legume interactions fix about half the amount of nitrogen (about 10-180 Kg hectare^{-1} year^{-1}) that the *Rhizobium*-legume interactions do. This is, however, still a greater amount than is fixed by free-living bacteria and is economically important especially in forestry situations.

Techniques for Bacterial Identification:

In population studies of bacterial species, including *Rhizobium*, it is often desirable to use identification methods which not only distinguish among strains but also provide information that can be used to study populations for classification and identification purposes. A wide variety of techniques have been employed to study *Rhizobium* populations. These methods, which are described below, have included serological methods, electrophoretic studies of cellular proteins, antibiotic resistance and molecular techniques such as observation of restriction digest patterns and use of gene probes. Several of these methods have been used as complements to other techniques in order to confirm observations. Differences between results obtained with the various techniques testify to the variation that exists between the *Rhizobium* species and highlight the need to use several assays and not to rely upon one. These differences also indicate some of the problems encountered when attempting to perform bacterial classification.

Intrinsic Antibiotic Resistance (IAR):

The ability of a bacterial strain to grow in the presence of an antibiotic indicates the presence, within the strain, of resistance genes. This resistance to antibiotics is known as the Intrinsic Antibiotic Resistance (IAR) of the isolate and can be employed to identify the bacterial strain. IAR profiles for individual isolates can be produced by scoring the ability of each isolate to grow in the presence of a variety of antibiotics. Each strain is then given a profile based on its resistance or sensitivity to each antibiotic. These profiles can be used for statistical analysis of the bacterial population.

IAR has been a method of bacterial strain identification favoured by some researchers and has been used on studies of diversity of *Rhizobium* species (Sinclair and Eaglesham, 1984; Glynn *et al.*, 1985; Turco and Bezdicek, 1987; Brockman and Bezdicek, 1989), *Bradyrhizobium* species (Meyer and Pueppke, 1980; Sadowsky *et al.*, 1987b), *Frankia* species (Carrasco *et al.*, 1995) and in tests on preference inheritance for strains of *R. leguminosarum* biovar *trifolii* by white clover (*Trifolium*

repens) (Hardarson and Jones, 1979). Mc Laughlin *et al.*, (1984) used IAR to classify natural *Rhizobium* populations in a field prior to the introduction of an inoculum strain. They found that the longest surviving inoculum strain had a similar IAR pattern to 50% of the natural isolates. The technique of IAR and its usefulness in identification of *Rhizobium* strains has been questioned by Stein *et al.*, (1982) and Chanway and Holl (1986). Both groups reported inconsistencies with IAR, even among different colonies of the same strain, although serological methods used for comparison allowed unambiguous strain identification. Stein *et al.*, (1982) have even reported that variations could be induced by changes in incubation time and inoculum concentration. Sadowsky *et al.*, (1987b) who studied the relationship between genetic diversity and nodulation ability in *B. japonicum* also reported finding no correlation between IAR and strain nodulation ability although SDS-PAGE and restriction digest patterns did produce a positive correlation with nodulation abilities. The use of IAR to divide strains between two serologically based clusters has been reported by Turco and Bezdicek, (1987) who suggested that antibiotic resistance could be the result of cell wall changes hence the serological differences. Brockman and Bezdicek (1989) also report a correlation between serogroup and IAR among strains of *R. leguminosarum* biovar *viciae*. Glynn *et al.*, (1985) report a correlation between IAR and both plasmid profiles and restriction digest patterns of strains of *R. leguminosarum* biovar *trifolii*. Bromfield *et al.*, (1982) also used IAR to differentiate strains of *Rhizobium*. They found that 15 of the 16 isolates of *R. leguminosarum* biovar *phaseoli* could be identified using this technique and that results correlated with fluorescent antibody tests. However, it was found that slow growing *Rhizobium* isolates could not be clearly differentiated using IAR screening. These observations appear to indicate that for certain applications IAR could provide a rapid and sensitive test for screening *Rhizobium* strains. However, for the purposes of identifying *Rhizobium* strains within a natural population it is not necessarily the best method to use. Chanway and Holl (1986) have reported that IAR results from mixed cultures were not totally reliable in identifying individual strains. They even report that, in some cases, strains revealed

IAR patterns which were composites of both inocula thus suggesting genetic exchange has occurred. If this is the result with only two strains the possible combinations within a soil community would be limitless thus rendering IAR screening useless.

One of the major problems with IAR is the need to culture each strain in order to ascertain its particular antibiotic resistance characteristics. In a mixed population this would involve slow, laborious work and may not even be possible. Given the alternative approaches, some of which are faster, more sensitive and reproducible, IAR does not appear to be the best method available for strain identification. One of the alternative methods for strain identification is the use of serological techniques.

Serology:

The use of various serological methods to tag outer membrane structures of bacterial cells with antibodies have been reported and developed over the past several years. A large, comprehensive review of the serological techniques applicable to plant bacterial studies, at the time, was carried out by Schaad (1979). This review included the techniques of agglutination and gel immunodiffusion, both of which result in the formation of a visible precipitate, the former in a solution and the later as a line on an agar plate. The use of tube agglutination and gel immunodiffusion has reported in the study of *R. leguminosarum* biovar *trifolii* (Dughri and Bottomley, 1983, 1984) and *R. meliloti* (Fuquay *et al.*, 1984). Dughri and Bottomley, (1983, 1984) categorised strains of *R. leguminosarum* biovar *trifolii*, isolated from nodules of subterranean clover (*Trifolium subterraneum* L.), into four serogroups while Fuquay *et al.*, (1984) identified five serogroups within isolates of *R. meliloti*. The serogroups, identified by Fuquay *et al.*, (1984), agreed with strain groupings based protein profiles.

The technique of Enzyme-Linked Immunosorbent Assay (ELISA) involves fixing the antigen to a plastic support (usually a micro-titre plate). A solution of antibodies is added to the well and antibodies attach to the antigen. This solution is removed and unbound antibodies are washed away before a second antibody, with an enzyme moiety, is added. This second antibody binds to the first antibody and any

excess is removed before the amount of bound enzyme-linked antibody is detected, usually by a colour reaction. Colour intensities can then be used to determine the original antigen concentration. ELISA has been used to study a variety of species of *Rhizobium* (Kishinevsky and Bar-Joseph, 1978; Berger *et al.*, 1979; Morley and Jones, 1980; Ahmad *et al.*, 1981; Jones and Morley, 1981; Martensson *et al.*, 1984; Wright *et al.*, 1986; Martensson *et al.*, 1987) and has been modified in order to improve specificity. Kishinevsky and Bar-Joseph (1978) used ELISA to confirm agglutination and immunodiffusion results from *Rhizobium* isolated from peanut plants (*Arachis hypogaea*). Improvements to the ELISA technique described by Kishinevsky and Bar-Joseph (1978) have been made by Berger *et al.*, (1979) who used a double antibody 'sandwich' by producing rabbit antibodies to the *Rhizobium* cells and then producing sheep antibodies to the rabbit antibodies. The enzyme, alkaline phosphate, was attached to the sheep antibodies. Berger *et al.*, (1979) used this method to identify *Rhizobium* cultures and isolates extracted from lentil nodules. Further improvements where made by Morley and Jones, (1980) and Martensson *et al.*, (1984) who provided a fluorescent substrate and changed the enzyme alkaline phosphatase for B-galactosidase respectively. Wright *et al.*, (1986) subsequently produced monoclonal antibodies to *Rhizobium leguminosarum* biovar *trifolii* by growing a cell culture of spleen cells from immunised mice. These antibodies were employed to identify an inoculum strain from the root nodules of clover grown in unsterilised soils. The results indicated that the use of monoclonal antibodies further improved the specificity of the ELISA technique.

A further serological technique reported as useful for *Rhizobium* identification utilises fluorescent antibodies. The technique has been used to study competition of nodulation among various strains of *R. leguminosarum* biovar *viciae* on lentils (May and Bohlool, 1983) and *R. japonicum* on soybeans (Moawad *et al.*, 1984). Robert and Schmidt, (1983) have reported the use of fluorescent antibodies in their work on the population changes and persistence of an introduced *R. leguminosarum* biovar *phaseoli* strain into soil containing a resident, natural population of such bacteria.

Fluorescent antibodies have also been used in conjunction with intrinsic antibiotic resistance to study possible factors which could effect the ability to recover certain strains of *R. leguminosarum* biovar *viciae* (Turco and Bezdicek, 1987).

Although of importance in the study of population dynamics, especially in relation to persistence of introduced inoculum strains into the soil, the use of fluorescent antibodies does have a major disadvantage; the need to culture the target strain in order that antibodies can be produced. The requirement to produce antibodies from a pure culture necessarily means that those *Rhizobium* strains for which no isolation technique has been developed cannot be identified. A further problem with this technique is the lack of specificity which can result in inaccurate results owing to cross-reactivity. For the purposes of single strain identification this is not acceptable. A report by Ames-Gottfred *et al.*, (1989) on the use of Chrome Azurol S agar plates to complement serological techniques may be of use in overcoming problems of cross-reactivity observed with Fluorescent antibodies.

The use of other techniques such as isoenzyme electrophoresis can also be used as a means of identifying individual strains and so overcome the problems faced when using serological methods. By elucidating several profiles on one gel it is also possible to provide comparisons between strains.

Isoenzyme electrophoresis:

The use of isoenzyme patterns for bacterial classification has been described for many bacterial species including *Escherichia coli* (Ochman *et al.*, 1983; Whittam *et al.*, 1983; Ochman and Selander, 1984), *Shigella* (Ochman *et al.*, 1983), *Bordetella* (Musser *et al.*, 1986) and *Penicillium* (Cruickshank and Pitt, 1987). Two reviews by Selander *et al.*, (1986) and Selander *et al.*, (1987) describe the methods employed in producing multi-locus enzyme electrophoretic patterns and also the main uses of these patterns for studies of bacterial population genetics.

Reports of electrophoretic separation of cellular proteins in *Rhizobium* for studies of enzymatic polymorphism can be found as early as 1969 (Fottrell and O'Hora,

1969). However, early researchers used paper electrophoresis and starch gel electrophoresis in the study of polymorphisms of only one type of enzyme such as 3-Hydroxybutyrate Dehyrogenase (Fottrell and O'Hora, 1969) or the esterases (Murphy and Masterson, 1970). In conjunction with the study of polymorphisms of more than one enzyme, polyacrylamide gel electrophoresis (PAGE) began to supersede the techniques of paper electrophoresis and starch gel electrophoresis, although Engvild and Neilsen (1985) used starch gels to study isoenzymes of five, mainly pea or vetch, *Rhizobium* enzymes. In that study they found it was possible to cluster the isolates into groups based on isoenzyme patterns. PAGE of *Rhizobium* isoenzymes has been described by Mytton *et al.*, (1977) and used for population studies on the species *R. leguminosarum* biovars *viciae* (Young, 1985; Young *et al.*, 1987; Segovia *et al.*, 1991), *trifolii* (Young, 1985; Harrison *et al.*, 1987, 1989a, b) and *phaseoli* (Young, 1985; Pinero *et al.*, 1988) and the species *R. meliloti* (Young, 1985). The study by Young, (1985) found that most *R. meliloti* isolates were distinct from the *R. leguminosarum* species. It also revealed a limited number of electrophoretic types (ETs) (three in *R. meliloti* and 14 *R. leguminosarum*) were present in all isolates and that some were present in only one biovar of the *R. leguminosarum* while others were shared between the three biovars, *trifolii*, *viciae* and *phaseoli*. The study by Young *et al.*, (1987) revealed the presence of variations in ETs between strains isolated from the lateral or primary roots of pea plants. This study also found a large diversity in ETs on each plant although overall ETs per plant were quite similar. Harrison *et al.*, (1989a) showed that there were few alleles per enzyme loci but that at the three sites examined differences in strain polymorphism could be detected. Harrison *et al.*, (1989b) employed ETs to examine host strain preference for a range of *R. leguminosarum* bv *trifolii* isolates. This study revealed no preference for a particular ET and also indicated that there was no change in the strain competitiveness following serial dilution of soil samples. Work by Pinero *et al.*, (1988) indicated that although there was little change at each enzyme loci overall the strains comprising *R. leguminosarum* bv *phaseoli* were not highly related but most likely comprised a mixed assemblage of

strains. They suggest that chromosomal variants, such as ETs, rather than plasmid borne variations would provide a more reliable method of taxonomically classifying the *Rhizobium* species.

The use of Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) in conjunction with other identification techniques (as described above) has been reported for studies on populations of *R. leguminosarum* and *R. meliloti*. Dughri and Bottomley, (1983, 1984) employed SDS-PAGE to elucidate protein profiles of *R. leguminosarum* bv *trifolii* strains previously grouped into serogroups based on the results of agglutination tests. They found that many strains of the same serogroup had similar protein patterns but that among strains of serogroup 27 (the dominant group in limed, acidic soils) there were four and five strains types in the two serotypes 27-A and 27-B respectively. Demezas *et al.*, (1991) found 16 ETs in a collection of strains from *R. leguminosarum* bvs *trifolii* and *viciae* and *R. meliloti*. They also found that RFLPs, produced using two chromosomal DNA probes, correlated with the ET findings. A study of *B. japonicum* isolates by Kamicker and Brill, (1986) indicated that ETs were more useful, than serotyping, for strain identification in this species. They also found that of the 543 nodule isolates examined none were from the two inoculum strains used but were all indigenous isolates. Sadowsky *et al.*, (1987b) used SDS-PAGE and other techniques to study 20 isolates of *B. japonicum* serogroup 123. In this study they found a correlation between SDS-PAGE groupings and the nodulation ability of the isolates which was confirmed with genomic DNA digestion patterns and RFLPs produced using a *nifHD* gene probe. Other groups have employed SDS-PAGE to examine *R. meliloti* isolates. In a study of 232 isolates Eardly *et al.*, (1990) identified 50 distinct ETs which were used to reveal the presence of two subgroups within the species *R. meliloti*. Fuquay *et al.*, (1984) used SDS-PAGE to assess protein profiles from *R. meliloti* isolates prior to creating antisera to these strains. They suggested that SDS-PAGE provided a useful method of screening a *Rhizobium* population to ensure that specific antisera production is representative of the whole population. A recent study by Delajudie *et al.*, (1994) of 80 isolates from *Sesbania* and *Acacia* growing in

Senegal (W. Africa) indicated that 50 strains could be clustered into three ETs whilst the remaining 30 strains could not be grouped. Two of the three ETs differed from groups containing reference strains, which had been employed to provide a comparison. Further investigations, using various other characteristics including DNA-DNA hybridisation and 16S rRNA sequencing, lead Delajudie *et al.* (1994) to propose amending the genus *Sinorhizobium* by adding three new species, one of which is presently classified as *R. meliloti*.

Although providing a powerful tool for the study of population genetics, gel electrophoresis of isoenzymes may give a biased description of natural bacterial populations. Estimations that only two-thirds of amino acid substitutions (or 25 percent of point mutations at the DNA level) result in a change in the net electrical charge within a protein have been reported by Ayala, (1982). This means that some variations among bacteria may go undetected by the use of the SDS-PAGE method. Identification techniques that act at the molecular DNA level should be more discriminating and therefore overcoming these problems.

Molecular techniques:

Plasmid profiling:

Molecular analytical techniques which have been used to study *Rhizobium* populations include plasmid profiling (Glynn *et al.*, 1985; Meade *et al.*, 1985; Bromfield *et al.*, 1987; Broughton *et al.*, 1987; Mozo *et al.*, 1988; Brockman and Bezdicek, 1989; Harrison *et al.*, 1989a, b; Hartmann and Amarger, 1991; Laguerre *et al.*, 1992a, b). One of the major problems encountered when trying to study the large plasmids (including the symbiotic [Sym] plasmids) of *Rhizobium* is that of breakage which is easily done when manipulating such large and delicate structures. For this reason manipulation must be kept to a minimum. The technique of direct lysis of the cells in the agarose gels, as first used by Eckhardt to analyse *Eschericia coli* plasmids

(Eckhardt, 1978), is now the most widely used method for obtaining plasmid profiles from *Rhizobium* strains.

Broughton *et al.*, (1987) used plasmid profiles to investigate the process of symbiotic plasmid transfer among *R. meliloti* strains in the rhizosphere of *Medicago sativa*. Plasmid profiles have been used to study population diversity among strains of *Rhizobium* isolated from a variety of origins (Glynn *et al.*, 1985; Mozo *et al.*, 1988; Harrison *et al.*, 1989a). Brockman and Bezdicek (1989) and Laguerre *et al.*, (1992a, b) also employed plasmid profiling to study variation in natural populations, however, they used isolates from a limited source. Meade *et al.*, (1985) and Harrison *et al.*, (1989b) used plasmid profiling in investigations of *Rhizobium* strain competition. Meade's group employed the technique for investigations on an inoculum strain whereas Harrison *et al.*, studied natural populations by searching for the predominance of one strain. Hartmann and Amarger, (1991) used plasmid profiles to compare a collection of nodulating *R. meliloti* isolates to a commercially used *R. meliloti* strain. Following confirmation of their observations with restriction digests and Insertion Sequence (IS) fingerprints they concluded that plasmid profiling was an acceptable method of assessing population diversity.

Plasmid profiles were employed by Burr *et al.*, (1995) to study *Agrobacterium vitis* isolates in decaying grape roots and canes. This group initially inoculated the groups with six strains, all of which had unique plasmid profiles. They examined the profiles of isolates recovered over a period of 23 months and found some isolates were recovered more frequently than others, but that each isolate was recovered at least twice in that time. Only 18 from 133 profile observed were not associated with the original inoculum strains.

Prior to this, plasmids were characterised on the basis of their size (Nuti *et al.*, 1977; Casse *et al.*, 1979; Spitzbarth *et al.*, 1979; Tichy and Lotz, 1981; Mink and Sik, 1983; Burkardt and Burkardt, 1984), sequence homology (Jouanin *et al.*, 1981) and various other traits which have been described in a review article (Denaire *et al.*,

1981). The use of DNA homology in conjunction with other typing techniques led Jarvis *et al.*, (1982) to propose that a new species, *R. loti*, should be formed.

DNA probes:

Further studies of DNA homology led to the development of species-specific and strain-specific probes which can be employed for identification of particular species or strains. Bjourson and Cooper, (1988) used strain-specific probes, isolated using two different subtraction-hybridisation techniques (mixed-phase [heterogeneous] and single-phase [homogeneous]), to distinguish among different isolates of *R. loti*. Symbiotic-specific DNA probes, have been used by Krol *et al.*, (1982), Watson and Schofield, (1985), Schofield *et al.*, (1987), Mozo *et al.*, (1988), Young and Wexler (1988) and Engvild *et al.*, (1990) to examine differences in the three biovars of *R. leguminosarum*. Hodgson and Roberts, (1983) also used strain-specific probes and colony hybridisation to identify *R. trifolii* strains occupying nodules on clover. Species-specific probes have also been used to evaluate *Rhizobium* from soil samples (Fredrickson *et al.*, 1988; Linne von Berg and Bothe, 1992) and from crushed root nodules (Cooper *et al.*, 1987). Holben *et al.*, (1988) used a DNA sequence, unique to a *B. japonicum* strain, as a probe to detect its presence in a soil sample. Fredrickson *et al.*, (1988) used the transposon Tn5 as a marker to trace *R. leguminosarum* biovar *viciae* and *Pseudomonas putida* in soil. They also employed other techniques such as plate counts and fluorescent antibody analysis in support and report comparable results for all methods used. DNA probes (*nif* and 16S rRNA genes) and other techniques have also been used to study the genus *Frankia*. These have been described in a review by Lechevalier (1994). The use of insertion sequences (IS) as probes has also been reported by Wheatcroft and Watson, (1987, 1988a, b), Hartmann and Amarger, (1991), Simon *et al.*, (1991) and Kosier *et al.*, (1993).

DNA hybridisation was also used by Laguerre *et al.*, (1992b), to study plasmid profiles and RFLP patterns using plasmid and chromosomal specific probes. Their results indicate that chromosomal probes alone could not be used for classification

purposes although symbiotic plasmid probes could. They also identified one group of strains, isolated from beans, which showed no similarity to the species *R. leguminosarum*. A second group they isolated displayed a greater affinity with *R. leguminosarum* biovars *trifolii* and *viciae* than with *R. leguminosarum* biovar *phaseoli*. They concluded that this could be due to the low inoculation specificity of the host plant, *Phaseolus vulgaris*.

DNA Fingerprinting Techniques:

Restriction Digest Patterns:

One of the simplest methods to produce DNA fingerprints is to digest the total genomic DNA with restriction enzymes. The cleaved DNA can then be separated by gel electrophoresis to produce a DNA profile which can be employed for classification purposes. Mielenz *et al.*, (1979) used this method to differentiate between strains of *R. trifolii*, *R. meliloti* and *R. japonicum*. This group also used the patterns to show that *R. trifolii* mutants, which nodulated soybean, where derivatives of a strain of *R. japonicum*. Schulz *et al.*, (1993) used the rare cutting enzymes XbaI, SfiI and SpeI to digest DNA from *A. vitis*. They separated the fragments using pulsed field gel electrophoresis (PFGE) and were able to classify the strains into six groups based on their restriction digest patterns. They also found a correlation between RFLP patterns, produced using IS, and each group identified from digest pattern analysis. DNA restriction patterns have also been employed, in conjunction with other techniques including IAR, SDS-PAGE and plasmid profiling to study *R. trifolii* (Glynn *et al.*, 1985) and *B. japonicum* (Sadowsky *et al.*, 1987b). Laguerre *et al.*, (1992a) also used DNA restriction patterns in conjunction with plasmid profiling and RFLPs to study field populations of *R. leguminosarum* biovar *viciae*. Hartmann and Amarger, (1991) used restriction digest patterns to confirm observations made from plasmid profile analysis.

DNA restriction patterns, although quite easily produced, are complex and not easily interpreted, especially at the extremes of the digested DNA size range. The use of labelled DNA probes to produce hybridisation patterns provides a better method of studying these digests as the final profile is much simpler and easier to interpret.

The production of these patterns from restriction digested DNA involves transferring the DNA to a solid support prior to hybridising it with a labelled DNA probe. The method of DNA transfer and detection as first described by Southern, (1975) is still used extensively with small modifications as necessary.

Restriction Fragment Length Polymorphism (RFLP):

Restriction Fragment Length Polymorphisms (RFLPs) can be used to study a wide variety of biological organisms including prokaryotic and eukaryotic species. Reports have been made of the use of RFLPs for studies of mammalian species including man (Botstein *et al.*, 1980) and horses (Hopkins *et al.*, 1991). RFLPs have also been used to study plant species including beans (*Phaseolus vulgaris*) (Nodari *et al.*, 1993a, b). Bacterial species have also been studied using RFLPs. This includes useful bacteria such as *Lactobacillus helveticus* (de los Reyes-Gavilan *et al.*, 1992) and plant pathogens such as *Xanthomonas campestris* (Gabriel *et al.*, 1988). RFLPs have also been employed in the study of most Rhizobial species including *B. japonicum* (Stanley *et al.*, 1985; Gibson *et al.*, 1990; Hartmann *et al.*, 1992; Vanberkum *et al.*, 1993; Madrzak *et al.*, 1995), *R. meliloti* (Hartmann and Amarger, 1991; Simon *et al.*, 1991; Kosier *et al.*, 1993; Brunel *et al.*, 1996; Paffetti *et al.*, 1996), *R. tropici* (Geniaux *et al.*, 1993), *R. galegae* (Nick and Lindstrom, 1994; Selenskapobell *et al.*, 1995), *R. leguminosarum* (Laguerre *et al.*, 1992b; Hirsch *et al.*, 1993), and the specific biovars, *R. leguminosarum* biovar *trifolii* (Demezas *et al.*, 1991, 1995), biovar *viciae* (Laguerre *et al.*, 1992a) and biovar *phaseoli* (Geniaux *et al.*, 1993).

Geniaux *et al.*, (1993) used several DNA probes, including a *nifH* probe, to generate RFLP patterns from strains of *Rhizobium* which had been isolated from bean nodules. Using these RFLP patterns, plasmid profiles and the ability, of the isolates, to

nodulate an alternative host (*Leucaena leucocephala*) they grouped the strains into three types, *R. leguminosarum* bv *phaseoli*, *R. tropici* or *R. etli* bv *phaseoli*. A larger study, by Laguerre *et al.*, (1992b), of isolates of *R. leguminosarum* using plasmid profiles and RFLPs revealed that grouping isolates into their respective biovars was only possible when using Sym plasmid probes. They also found, using chromosomal probes, that many (40%) of the isolates had similar chromosomal RFLP profiles. Demezas *et al.*, (1991) also employed chromosomal probes to study a selection of isolates from *R. leguminosarum* bv *trifolii*, bv *viciae* and *R. meliloti*. This study revealed 16 RFLP patterns within the strains which corresponded to ET patterns obtained previously for these isolates. From the RFLPs it was possible to calculate genetic relationships among the isolates, however, when Sym plasmid DNA probes were used to create RFLP patterns this was not possible. However, a study by Vanberkum *et al.*, (1993), indicated that RFLP patterns produced using probe pRJUT10, a nodulation locus, could be employed to cluster isolates of *B. japonicum*, serogroup 110, into four groups which showed a degree of similarity of 70 % or less. Madrzak *et al.*, (1995) also used Sym plasmid DNA probes (*nif* and *nod* gene probes) to examine RFLP profiles from populations of *B. japonicum* from Polish soils. In that study they identified two major strain groups within these isolates and confirmed these observations with protein profile analysis, serological reactions and REP-PCR profiles. A further investigation by Stanley *et al.*, (1985) also revealed the existence of two major sub-groups within *R. japonicum* (now *B. japonicum*) and that one of these also displayed the presence of two sub-lines. A two year study by Demezas *et al.*, (1995) found that the number of Sym plasmid type strains (based on RFLPs produced using Sym plasmid DNA probes) changed from one year to the next while no difference in the number of chromosomal type strains was observed. This indicated a larger degree of variation occurred within Sym plasmid DNA than occurred within chromosomal DNA. A similar study was conducted by Gibson *et al.*, (1990). They used RFLPs to study the genetic instability of *B. japonicum* strain CB1809, a widely used Australian

inoculum, in soils over a period of up to nine years. However this study revealed no change in RFLP patterns although serological variations were observed.

The use of RFLP analysis to study ineffective rhizobia (S isolates) which nodulate white clover plants growing on contaminated soils was reported by Hirsch *et al.*, (1993). In this study they found that all isolates had similar RFLP patterns which indicated a common ancestor, however, the isolates were not related to those strains found on white clover growing in clean soils. This observation was confirmed by inoculating other clover species with the isolates. This revealed that effective nodules were produced on subterranean clover only.

The type of probe employed, to produce RFLP profiles, varies and depends upon the particular interests of the research group. The use of insertion sequences (IS) for RFLP production was reported by Hartmann and Amarger (1991) who employed ISRm1 as a probe to study a representative sample (32 from 125 isolates) of *R. meliloti* isolates. They found that IS fingerprints revealed 16 strain types which correlated well with plasmid profile analysis. IS have also been used as probes in studies of *R. meliloti* by Wheatcroft and Watson, (1987), Simon *et al.*, (1991) and Kosier *et al.*, (1993). Repeat sequences (RS) have also been employed, as probes, in studies of *Giardia intestinalis*, a parasitic protozoal species, (Ey *et al.*, 1992), *R. leguminosarum* biovar *trifolii* (Watson and Schofield, 1985; Schofield *et al.*, 1987; Harrison *et al.*, 1988) and *B. japonicum* (Gibson *et al.*, 1990; Hartmann *et al.*, 1992; Minamisawa *et al.*, 1992). The use of repeat sequences (RS) to produce RFLPs was reported by Hartmann *et al.*, (1992) who employed these to differentiated individual strains of the same serogroup of *B. japonicum*. Nick and Lindstrom, (1994) and Selenskapobell *et al.*, (1995) also used RS to study *R. galegae*, however, they used RS-PCR to study the bacterial strains and confirmed these results with RFLP analysis.

Chromosomal gene probes have also been employed to create RFLP patterns from Rhizobial species (Laguerre *et al.*, 1992a, b; Hirsch *et al.*, 1993; Demezas *et al.*, 1991, 1995) although all three groups also used plasmid derived probes for comparison purposes. Plasmid derived probes, such as the *nif* and *nod* gene clusters,

comparison purposes. Plasmid derived probes, such as the *nif* and *nod* gene clusters, have also been used to study Rhizobial isolates (Christensen and Schubert, 1983; Sadowsky *et al.*, 1987a; Mozo *et al.*, 1988; Segovia *et al.*, 1991; Geniaux *et al.*, 1993; Madrzak *et al.*, 1995). Kuykendall *et al.*, (1996) generated RFLP profiles for *Leucaena*-nodulating *Rhizobium* from a selection of six random cosmid clones. Lee *et al.*, (1992) used 18 different probes to produce RFLPs from Mycoplasma like organisms (MLO). Data from the 18 probes were amalgamated into a single average similarity matrix using the Jaccard (Jaccard, 1901) coefficient of similarity. This matrix was then used to find relatedness among three MLO strain clusters associated with Canadian Peach (Eastern) X Disease, Western X Disease and Clover Yellow Edge.

Although of value for strain identification and classification, RFLPs are of limited use owing to the need to provide pure DNA samples. The use of specific probes also results in an examination of a small part of the total genome (i.e. that part from which the probe was derived) to the exclusion of other areas which may have a greater variation. A technique of DNA examination which considers the whole genome may be of more use for phylogenetic studies. The use of DNA amplification to create fingerprints provides such a method. Harrison *et al.*, (1992) have reported that it is possible to amplify *Rhizobium* DNA from crushed roots so this method also eliminates the need to provide pure DNA samples.

DNA amplification:

The polymerase chain reaction (PCR) as described by Mullis *et al.*, (1986) is a method of generating large quantities of DNA from a very small sample comprising only a few DNA molecules. The method involves mixing primers, which can be either oligonucleotides or short sequences of DNA of known sequence, with the basic DNA components and the template DNA to be amplified. The primers are selected for their complementation to known gene sequences, or sections of DNA of interest, within the template DNA. The enzyme *Taq* polymerase, from *Thermus aquaticus*, a bacterial species found in hot springs, is used to amplify the DNA. The standard DNA

amplification reaction can only provide information regarding the presence or absence of a particular DNA region of interest within the genome. Bereswill *et al.*, (1992) used this feature of the reaction to develop a detection method for *Erwinia amylovora*, a pathogen which causes fireblight on fruit trees. By using two primers, derived from a 29 Kb plasmid present in all *E. amylovora* isolates, they could detect the pathogen by the presence of a 0.9 Kb band amplified by these two primers. However, the standard PCR reaction, with the exception of microsatellites, has only a limited use for phylogenetic studies owing to the limited number of amplification products (usually one band) which are produced. Microsatellites, which are produced using a standard PCR type reaction, can be used for parentage testing and phylogenetic studies. These tests involve examining variations in the number of dimer repeat sequences within hyper-variable regions from two individuals. These variable regions are initially amplified using two primers which bound these regions. With slight modifications to the reaction it can be made to provide data which is of use for identification and characterisation purposes.

DNA amplification and sequencing:

DNA amplification followed by sequencing of the 16S Ribosomal gene (16S rRNA) fragment has allowed phylogenetic studies of *Rhizobium* strains to be made. Using this technique Segovia *et al.*, (1991) and Young *et al.*, (1991) have examined relationships between *Rhizobium* strains. Jarvis *et al.*, (1992) employed this method to compared DNA sequence data from fast growing strains (isolated from soybean) with sequence data from other strains of *Rhizobium*, *Bradyrhizobium* and *Agrobacterium*. They concluded that these fast growers had a greater homology with the *Rhizobium* species than with the other species and proposed they be called *R. fredii*. Eardly *et al.*, (1992) employed similar techniques to classify isolates known to nodulate both alfalfa (*Medicago sativa*, L.) and beans (*Phaseolus vulgaris*, L. Savi). They found that those strains nodulating both alfalfa and beans resembled strains isolated from *P. vulgaris* in phenotypic characteristics. However, when comparisons of

16S rRNA gene sequences were made the dual nodulating strains formed a new group which contains strains previously known as *R. leguminosarum* biovar *phaseoli* type I.

Sawada *et al.*, (1993) used DNA amplification and sequencing of the 16S rRNA genes to study *Agrobacterium*, *Rhizobium*, *Azorhizobium* and *Bradyrhizobium*. They found it impossible to differentiate between *Agrobacterium* and *Rhizobium* isolates which formed a separate cluster from the remaining isolates. However, they could differentiate between the biovars (bv I, bv II, bv *rubi* and bv *vitis*) of *Agrobacterium*. These observations were confirmed by Willems and Collins (1993) who found *Agrobacterium* and *Rhizobium* formed a single cluster, based on 16S rRNA sequences, which was distinct from clusters formed by isolates of *Azorhizobium* and *Bradyrhizobium*.

DNA amplification and sequencing was employed by Cournoyer *et al.*, (1993) to study *Frankia*, who found two major phylogenetic clusters based on 16S rRNA and *nif*H-D intergene sequences. These groups corresponded to previously found groupings based on infection mechanisms of *Frankia* isolates. From an analysis of 16S rRNA sequences of isolates of *Blastobacter*, Hugenholtz *et al.*, (1994) found that two of the culturable species, *B. aggregatus* and *B. capsulatus*, were very similar to *A. tumifaciens* whilst isolates of *B. denitrificans* grouped with isolates of *B. japonicum*. This suggested that the taxonomy of *Blastobacter* required revision as it appeared to be composed of several different genera. A revision of the family Frankiaceae, based on 16S rRNA sequences, has been proposed by Normand *et al.*, (1996) who found four subdivisions within the genus *Frankia* but could find no relationship between the two genera *Frankia* and *Geodermatophilus*.

This technique is, however, very time-consuming and laborious owing to the need to sequence a large number of amplified products. A method which could provide as much data whilst also reducing time and labour would be advantageous.

RFLP-PCR:

A modification to the technique of 16S rRNA gene amplification and sequencing is to amplify the DNA but instead of sequencing the amplified products they are digested with restriction enzymes to produce RFLP type fingerprints. This technique, known as either RFLP-PCR or ARDRA (amplified ribosomal DNA-restriction analysis) was used by Vaneechoutte *et al.*, (1992) to distinguish 9 of the 13 taxa of the Comamonadaceae bacteria. Amarger *et al.*, (1994) also used RFLP-PCR of the 16S rRNA gene to classify *Rhizobium* which nodulated *Phaseolus vulgaris* L. into the species *R. leguminosarum* biovar *phaseoli* or *R. tropici*. They confirmed these observations using plasmid profiling and by assessing the ability of the isolates to nodulate *L. leucocephela*. The technique has been employed on studies of many of the Rhizobial species. Laguerre *et al.*, (1994) studied RFLP-PCR variations of the 16S rRNA genes in a representative sample of all the *Rhizobium* and *Bradyrhizobium* species. Results obtained using this method were in accordance with recognised taxonomic classifications. A study of *R. leguminosarum* biovars *trifolii*, *viciae*, and *phaseoli*, conducted by Laguerre *et al.*, (1996), showed that RFLP-PCR profiles from the *nod*D gene regions were biovar or species-specific whilst patterns from the intergeneric 16S-23S rRNA genes and the *nif*D-*nif*K genes regions were universal for the *Rhizobium* species. Using the inter 16S-23S rRNA gene regions and inter *nif*D-*nif*K gene regions Brunel *et al.*, (1996), found that *R. meliloti* fell into two separate groups. This observation is similar to those made by other groups using different assessment methods (Young, 1985; Eardly *et al.*, 1990; Dooley *et al.*, 1993 [work from this publication forms part of this thesis]). Analysis of the *nif*HD region of a spontaneous mutant of *Frankia* permitted Cournoyer and Normand, (1994) to classify the mutant as *F. alni*. Using RFLP-PCR of the *gln*II gene also allowed them to identify the parent strain of this mutant. RFLP-PCR was also employed by Jamann *et al.*, (1993), to study *Frankia* strains. They found classification based on RFLP-PCR correlated well with established taxonomic groups. Similar observation were made by Bosco *et al.*, (1996) who used RFLP-PCR to compare *Frankia* isolates from *Dryas*

drummondii to *Frankia* isolates from *Alnus*, *Causuarnia* and *Elaeagnus*. They found that the isolates from the *Dryas* did not cluster with the two established groups, *Alnus-Causuarnia* infective and *Elaeagnus* infective, but appeared distantly related to them.

RFLP-PCR profile analysis appears to be of use in classifying and identifying bacterial strains including *Rhizobium*, however, it relies upon variations within a small region of the total genomic DNA and so may not be representative of total variations between the isolates.

Repeat Sequences:

A second method of fingerprinting bacterial strains using DNA amplification is that of RS-PCR. Repeat sequences, as described above, were found to contain conserved inverted repeat sequences within their structure. Using these inverted repeats as primers it is possible to produce amplified DNA fingerprints with which bacterial strains can be identified or classified. These repeats were first reported in *R. meliloti* by De Bruijn (1992), who showed that they could be used to classify the *R. meliloti* strains. RS-PCR (REP-PCR/ERIC-PCR) has subsequently been used in studies of *R. leguminosarum* (Labes *et al.*, 1996; Laguerre *et al.*, 1996), *B. japonicum* (Judd *et al.*, 1993; Madrzak *et al.*, 1995), *R. loti* (Schneider and De Bruijn, 1996) and *R. galegae* (Nick and Lindstrom, 1994; Selenskapobell *et al.*, 1995).

RS-PCR appears to be of use for identification and classification of *Rhizobium* isolates. However, the method is limited by the use of a specific-primer (one which complements the RS) and by the relatively large size of the primers, 18 bp (REP sequence) and 22 bp (ERIC sequence) (De Bruijn, 1992). A method which employed random primers would result in an analysis which considered the whole genome and not just a select part of it. Smaller primers would also increase the number of primer binding sites within the genome which increases the chance of variation being detected.

RAPD fingerprinting:

One of the more useful (with regard to taxonomic studies) modifications to the PCR technique was reported by Williams *et al.*, (1990) and Welsh and McClelland, (1990). Both these groups employed random primers which do not complement known sequences within the template DNA. The resulting Randomly Amplified Polymorphic DNA (RAPD) profiles (or fingerprints) can be used for taxonomic classification at the species or strain level.

Williams *et al.*, (1990) used arbitrary primers of varying length to produce RAPD profiles which could be used to identify a number of individuals from differing eukaryotic and prokaryotic species. Welsh and McClelland, (1990) and Sellstedt *et al.*, (1992) used the technique to identify specific bacterial strains. Irelan and Meredith, (1996), used RAPDs to study the variation between and within the three biovars of *Agrobacterium*; *A. tumifaciens* bv I and bv II and *A. vitis* (bv III). The use of RAPDs for identification purposes has also been reported for fungal strains (Goodwin and Annis, 1991), plants (*Glycine max*) (Caetano-Anolles *et al.*, 1991), insects (Wilkerson *et al.*, 1993) and domestic animals (cattle, chickens, dogs, sheep and horses) (Bailey and Lear, 1994; Cushwa and Medrano, 1996).

Harrison *et al.*, (1992) reported the use of random arbitrary primers for the production of amplified DNA profiles from *R. leguminosarum* biovar *trifolii*. That study employed methods which produced direct DNA amplification from cell cultures and nodule tissue and indicated DNA conservation among certain isolates.

Welsh and McClelland, (1991) have also reported that different patterns were produced when two, different, random primers were used to prime the same amplification reaction. These fingerprints differed from those produced by the two primers when either one of them was used alone.

Fekete *et al.*, (1992) used five random primers, in single primer reactions, to produce RAPD profiles from *Brucella* strains. Using the Jaccard coefficient they produced similarity matrices for each of the primers. A single average matrix was constructed from the five individual matrices. This was then used to assess levels of

similarity among the strains. This technique is similar to that used by Lee *et al.*, (1992) for analysing RFLP data as described above.

Although RAPD analysis of *Rhizobium* isolates provides a quick and easy way to identify and classify the strains there are some problems which can be encountered. A report by Coutinho *et al.*, (1993) indicated that the age of the culture affected the RAPD profile produced and that, in order to produce reliable patterns, it was necessary to use young, fresh cultures. A further problem which can be faced is that of humic acid and other phenolic compounds which can be present in DNA extractions from environmental samples, in particular soil samples. There are a few reports on the affects of these compounds on the action of the enzyme *Taq* polymerase during DNA amplification, however these reports are contradictory. Tsai and Olsen (1992) do not report problems when they amplified 16S rRNA gene DNA from crude DNA extracts from soil and sediment samples. This would suggest *Taq* polymerase is unaffected by humics. Picard *et al.*, (1992) found that they needed to pass DNA extracted from soil through purification columns at least three times in order to clean it sufficiently for amplification. They also adjusted the conditions for DNA amplification. Bruce *et al.*, (1992) also cleaned DNA to remove brown residues, which indicate the presence of humic acid contaminates, before attempting amplification. Young *et al.*, (1993) suggested the technique of adding polyvinylpyrrolidone to agarose gels as a method of cleaning DNA samples sufficiently to allow DNA amplification to proceed.

Statistical analysis of DNA Fingerprints:

The use of mathematical analysis to interpret DNA fingerprints has been reported for restriction digest patterns (Nei and Li, 1979) and RFLPs (Castagna *et al.*, 1994; Laroche *et al.*, 1995; O'Donoughue *et al.*, 1994; Szalanski *et al.*, 1996; Sovinski *et al.*, 1996;) and RAPDs (Fujimori and Okuda, 1994; Kambhampati *et al.*, 1992; Okuda *et al.*, 1995; Oxelman, 1996; Paffetti *et al.*, 1996; Shi *et al.*, 1996; Stewart and Excoffier, 1996; Stewart *et al.*, 1996; Tanaka *et al.*, 1994). Various methods of analysis have been used on the different fingerprinting techniques with the method of

choice for each research group being based on personal preference. Nei's genetic distance (Nei and Li, 1979; Vancoppencolle *et al.*, 1995; Szalanski *et al.*, 1996) and the Hardy-Weinberg (Sovinski *et al.*, 1996) constant have been widely used methods for RFLP analysis although two recent papers by Castagna *et al.*, (1994) and O'Donoughue *et al.*, (1994) describe the use of Principal Coordinate Analysis (PCO) (Gower, 1966) for RFLP analysis. The interpretation of RAPD fingerprints from a small sample of isolates has followed along similar lines to that of RFLP analysis. These methods of comparison are relatively easy to perform on a few isolates, however, when larger numbers of strains are used an improved method is needed.

Multivariate statistical analysis provides a method of analysing large numbers of variables which univariate statistical methods are incapable of handling. An extensive summary by James and McCulloch, (1990) describes the main multivariate analysis techniques used in ecological and systematic studies. This summary includes a description of PCO and Cluster Analysis, both of which can be used to display relationships between organisms. PCO provides a method of displaying a complex relationship, between individuals, on a two or three dimensional graphical plot. However, relative spatial positions of individuals on the plot cannot be used to describe their relatedness but act only as a guide to this. Cluster Analysis on the other hand links individuals one-by-one depending on their level of similarity. The two closest related individuals are joined first at the highest level of similarity. As the level of similarity is decreased new isolates are introduced to the existing cluster or a new group is formed. This process continues until only one large amalgamation exists.

The use of PCO and Cluster Analysis to study biological systems has been reported for protein sequences (Higgins, 1992), soils (Oliver and Webster, 1989) and protein profiles from wheat (Bietz and Simpson, 1992). Vauterin and Vauterin, (1992) used Cluster Analysis and computer-aided comparisons to identify micro-organisms from their protein profiles. PCO and Cluster Analysis have also been employed in the analysis of RAPD data (Dooley *et al.*, 1993 [this publication forms part of this thesis]; Dye *et al.*, 1995). There have also been reports of the use of Cluster Analysis with

UPGMA (Unweighted Pair Group Method with Averages) (Haig *et al.*, 1994), Jaccard matching (Morgan *et al.*, 1995) and nearest neighbour grouping (Kambhampati *et al.*, 1992) for analysis of RAPD data. More recently RAPD profiles have been analysed using the techniques of Phylogenetic Analysis Using Parsimony (PAUP) (Fujimori and Okura, 1994; Tanaka *et al.*, 1994) and Analysis of MOlecular VAriance (AMOVA) (Excoffier *et al.*, 1992; Haig *et al.*, 1994; Peakall *et al.*, 1995; Paffetti *et al.*, 1996; Stewart and Excoffier, 1996; Stewart *et al.*, 1996) and HOmogeneity of MOlecular VAriance (HOMOVA) (Stewart and Excoffier, 1996; Stewart *et al.*, 1996).

The statistical techniques described produce taxonomic profiles which are in accordance with previously identified classifications. It therefore appears that for analysis of RAPD fingerprints it does not matter which of these methods is used as they all produce equally useful results for classifying *Rhizobium*.

Rhizobium Taxonomy:

The use of biochemical and molecular techniques has helped with the reclassification of the *Rhizobium*. However, owing to an incomplete, but ongoing, study of the *Rhizobium* and *Bradyrhizobium* genera the present taxonomic classification is not finalised. The most recent review of *Rhizobium* and *Bradyrhizobium* taxonomy was conducted by Martínez-Romero, (1994). The results of this review are summarised in Table 1.1.

The previous review of the Rhizobiaceae had been conducted by Elkan, (1992) and it varied from the one by Martínez-Romero in that those strains now classified as *R. fredii* and *R. xinjiangensis* had been classified *Sinorhizobium fredii* and *S. xinjiangensis* respectively. Also in the former review there was only one recognised species of *Bradyrhizobium*, *B. japonicum*. The work described in this report has been based on the classification system reviewed by Elkan, (1992). It was noted, however, that changes in the later review by Martínez-Romero, (1994) did not effect strains employed for this work.

Martínez-Romero *et al.*, (1991) have described a new species, *R. tropici*, (which was described by Elkan, 1992) which contains those strains previously known as *R. leguminosarum* bv *phaseoli* type II. Although I have noticed that within the group of strains of biovar *phaseoli* used in this study there is a certain degree of diversity which may be accountable by this new species I have not identified which groups individual strains belong to. These might be *R. leguminosarum* bv *phaseoli* type I or *R. tropici*.

The changes observed in strain classification over just the two year period between the reports made by Elkan and Martínez-Romero highlight the need for improved methods of rhizobia strain classification. In the period since the review by Martínez-Romero there have been further suggestions for classification changes. Delajudie *et al.*, (1994) suggest maintaining the genus *Sinorhizobium*. They even add two new species, *S. saheli* and *S. teranga* and move *R. meliloti* into this genus as a species, *S. meliloti*. There have also been proposals for new species including *B. liaoningense* (Xu *et al.*, 1995), *R. mediterraneum* (Nour *et al.*, 1995) and *R. tianshanense* (Chen *et al.*, 1995). Novikova *et al.*, (1994) also identified a group of *Rhizobium* strains, isolated from temperate regions, which formed a cluster separated from clusters formed of *Rhizobium*, *Bradyrhizobium* and *Agrobacterium* isolates.

With these fluxes in taxonomic groupings within the Rhizobiaceae family a study which could provide information about *Rhizobium*, based at the molecular level, may be useful for identification or classification and may also be of use for confirming previous observations. An investigation by Harrison *et al.*, (1992) indicated that RAPD fingerprinting provided a method of identifying *Rhizobium* strains from a small collection (12 isolates) of *R. leguminosarum* bv *trifolii* isolates. A study which examined a larger selection of isolates may provide more information on *Rhizobium* taxonomy. The study presented here employed a more diverse selection of isolates than that used by Harrison *et al.*, (1992) and was conducted with the intention of addressing the following aims:

Objectives:

1. To assess the use of RAPD fingerprinting for the classification and identification of *Rhizobium*, and in particular *R. leguminosarum* bv *trifolii*, isolates. This study also included the development of a suitable method of statistical analysis.

2. To identify and quantify species- and strain-specific DNA probes, arising from DNA amplification, which could be used for further studies of *Rhizobium* identification.

3. To derive genetic information regarding interspecies relationships.

4. To compare the results of these three main objectives with previously observed results.

The results of the preliminary investigation by Harrison *et al.* (1992) provided information which allowed the main aims of this study to be addressed. By using a wider selection of isolates of *Rhizobium* and *Bradyrhizobium* species and biovars it should be possible to assess the use of RAPD fingerprinting for rhizobial classification. A large selection of isolates should also provided sufficient data for performing a comprehensive investigation of suitable statistical analysis methods. By using isolates from a wider variety of species and biovars the identification of potential RAPD-derived species-specific DNA probes should be made easier. A larger number of overall strains would also increase the probability identifying strain-specific DNA probes.

The report presented here describes the results of an investigation conducted with the aims of addressing the four points detailed above.

Tables and Figures

Table 1.1: Taxonomy of the Rhizobiaceae family (from Martínez-Romero, 1994).

Recognised genera	Recognised species
Rhizobium	
	R. etli
	R. fredii
	R. galegae
	R. huakuii
	R. leguminosarum
	bv *viciae*
	bv *trifolii*
	bv *phaseoli*
	R. loti
	R. meliloti
	R. tropici
	R. xinjiangensis
Bradyrhizobium	
	B. elkanii
	B. japonicum
Azorhizobium	
	A. caulinodans

General Materials and Methods

General Materials and Methods

***Rhizobium* cultures and growth media.**

A selection of 84 *Rhizobium* and *Bradyrhizobium* strains were obtained from the AFRC Institute of Grassland and Environmental Research, Welsh Plant Breeding Station (IGER-WPBS), Aberystwyth, Dyfed. This collection contained 18 isolates each of *Rhizobium leguminosarum* biovar *trifolii*, *R. leguminosarum* biovar *viciae* and *R. leguminosarum* biovar *phaseoli* as well as 17 strains of *R. meliloti* and 13 *Bradyrhizobium* isolates from diverse geographical locations. These strains are shown in Table 2.1. The freeze dried cultures were reconstituted by opening the vials and adding 0.5 ml of PA solution (Table 2.2) (Hirsch *et al.*, 1984) to the culture. A sample (50 µl) of the *Rhizobium* suspension was plated onto TY agar plates (Table 2.2) (Beringer, 1974). Following two days incubation at 27°C pure colonies were streaked, for single colonies, onto a second TY plate. Pure colonies were again streaked for single colonies onto a third TY agar plate. *Rhizobium* cultures from these plates were stored in glycerol (20% v/v) at -20°C. These stored cultures formed the basis of this study.

Strains used in this work have been named, as far as possible, in accordance with the recognised classification described by Elkan (1992) (see Chapter 1). However, it is not known whether the isolates of *R. leguminosarum* biovar *phaseoli* used here constitute type I or *R. tropici* (Martinez-Romero *et al.*, 1991) (formerly known as type II) strains. For the purposes of this study they have been referred to as *R. leguminosarum* biovar *phaseoli*. It should also be noted that *Bradyrhizobium* strains isolated from Japonicum have been referred to as *B. japonicum*. All other *Bradyrhizobium* strains are referred to as *Bradyrhizobium* strains. This includes those strains isolated from lupins and previously referred to in the Rothamsted Collection of *Rhizobium* (RCR) catalogue as *B. lupini*.

General growth media employed in this study are described in Table 2.2 together with relevant references.

Preparation of bacterial cells.

Fresh cell cultures were produced, when required, by inoculating a sample (10 µl) of the stored culture from above into 10 ml of PA solution. Cells were grown overnight at 27° C in an orbital shaker to a concentration of approximately 10^8 cells ml^{-1}. Larger volumes of cells were produced, when necessary, by inoculating a greater volume of growth media with 1 ml of overnight culture. These cultures where stored for short periods (up to one week) at 4° C.

Rhizobium cultures were also maintained on TY agar plates for up to 2 weeks at 4° C and used as required. These plates were initially produced from the frozen cultures by spreading a 10 µl sample on to a plate and then streaking for single colonies following growth after 2 days. Frozen cultures were also screened to ensure purity using this method.

Preparation of bacterial cells for direct DNA amplification:

Cells were grown overnight at 27° C in PA medium to a concentration of approximately 10^8 cells ml^{-1}. A sample (0.5 ml) of each culture was centrifuged at 11600g and the supernatant removed. The resulting pellet was washed in 0.5 ml of sterile distilled water (SDW). Following a further centrifugation, the washed pellet was resuspended in SDW and further diluted to achieve a concentration of 10^4 to 10^5 cells per 50 µl reaction.

Conditions for DNA amplification:

Amplification reactions were carried out in 50 µl volumes. The reaction buffer contained 10 mM Tris (pH 8.3); 50 mM KCl; 2 mM $MgCl_2$; 0.001% gelatin. dNTPs were added to a concentration of 200 µM. Individual primers, whether used in single primer or double primer reactions, were added to their relevant concentrations (see

specific chapters). Taq polymerase (1 unit) (Advanced Biotechnologies Ltd, Leatherhead, Surrey, U.K.) was used to amplify DNA. All reactions were sealed with sterile mineral oil (Sigma Chemical Co., Poole, Dorset, U.K.) and a Hybaid Thermal Reactor (Hybaid Ltd., Teddington, Middlesex, U.K.) was used to apply the following temperature cycle: 3 min at 94° C, initial denaturation; 40 cycles at 94° C (1 min), 36° C (1 min), 72° C (2 min); 5 min at 72° C to provide a final extension period.

Statistical analysis of RAPD profiles:

RAPD product sizes where calculated by comparing the migration distance of each band against the distance migrated by the standard size markers. Each strain was then assigned a score on the presence or absence of each product size. A '1' was scored for the presence of a particular amplified band and a '0' for the absence of a band. Scores were analysed using GENSTAT (version 5, 1.3) to produce a similarity matrix. The program used to perform the analysis is shown in Appendix 1.

Construction of similarity matrices:

The similarity matrix displays, in a tabular-like form, the percentage similarities between strains thus permitting a comparison among strains to be made. Several different similarity matrices can be constructed from the same data depending on the matching coefficient employed.

Similarity matrices were constructed from the data using the simple matching coefficient. This method of matching considers all data points and gives equal weighting to all positions (loci). In the case of presence/absence (1/0) data this means all positions are considered, even if both strains show the absence of a particular band. Other matching coefficients, such as the Jaccard coefficient, would omit this band position from consideration. The rationale for simple matching can be summarised as follows:

$$Xi = Xj = 1 \text{ then } 1$$

$$Xi = Xj = 0 \text{ then } 1;$$

where X is the band under consideration in the two strains i and j.

In this case all band positions are considered and all contribute to the final percentage similarity. The presence of a band in both strains therefore returns an exact match, i.e. 100% similarity at that locus. Equally the absence of a band from both isolates produces a similarity of 100% at that locus. However, the presence of a band in one strain whilst it is absent from the other returns a similarity of 0% at that locus. The overall similarity between the two isolates is therefore equivalent to the average similarity between all the loci considered.

Principle Coordinate Analysis (PCO) was applied to the resultant similarity matrix and the first two ordinates from this plotted. This technique involves plotting all the data on a multi-dimensional graph with axis numbered 1 to *n*, where *n* is the number of samples examined. In the case of the samples in this study that means there are 84 axis, all of which meet at the origin (0,0) and all of which are perpendicular to each other! Obviously such a plot would be impossible to visualise on paper and so only the first two ordinates are drawn. This is performed by examining the whole multi-dimensional plot and extracting that axis which explains the greatest variation between the samples. This is known as ordinate 1. This procedure is repeated *n*-1 times, i.e. until all the ordinates are extracted. Each subsequent ordinate explains less information than the last. Only the first few ordinates are of value as they explain the greatest amount of variation. These ordinates can then be plotted on a two-dimensional (or possibly three-dimensional) graph with relative degrees of variation radiating out from the origin (0,0).

Cluster Analysis, using an average linking method, was also applied to the data. The results of this were added to the PCO plot in the form of shaded groups at three levels of similarity, 70-75%, 75-80% and 80-100%.

Total genomic DNA extraction:

Total genomic DNA was obtained from bacterial cells using the following method. Cells were grown as described above. The cell suspension (10 ml) was pelleted in a 5 ml tube at 11600g. The pellet was resuspended in 1 ml of TE, pH 7.6 (10 mM Tris-Cl, pH 7.6, 1 mM EDTA, pH 8.0) and 1.5 ml of TE, 2% w/v SDS (as TE with 2% w/v SDS) was added. The proteolytic enzyme Pronase E (Sigma) was added, to this suspension, to a final concentration of 1 mg ml^{-1} and the volume brought to 3 ml with TE. Following mixing by inversion, the tubes were incubated at 37° C for 30-60 min until the lysate solution turned clear. The solution was vortexed and DNA precipitated overnight at -20° C in 2.5 vols. of ethanol. The DNA was pelleted by centrifugation at 11600g for 5 min and the supernatant discarded. After drying, the DNA was resuspended in 500 µl of TE. RNAase (Ribonuclease A, Sigma) (from a stock solution of 10 mg/ml [Sambrook, 1989]) was added to a final concentration of 0.1 mg ml^{-1} and incubated at 37° C for 30 min. The protein elements within the solution were removed by two phenol-chloroform extractions (Sambrook, 1989) and a final chloroform extraction. The DNA was again precipitated overnight in ethanol at -20° C and recovered by centrifugation. After drying, the DNA was redissolved in 100 µl of TE and a sample (5 µl) run on a gel (as described below) in order to ascertain its concentration.

Conditions for Restriction Digests:

Total genomic DNA was digested using the restriction enzyme *Eco*R1 from the SuRE/Cut range of enzymes (Boehringer Mannheim, Bell Lane, Lewes, E. Sussex, U.K.). A sample of DNA (10 µg) was added to 2 µl of SuRE/Cut reaction buffer H, 1 Unit of enzyme was added and the volume brought to 20 µl using SDW. Digestion proceeded overnight at 37° C and was terminated by the addition of 5 µl of gel loading buffer (0.25% w/v Bromophenol blue, 40% v/v glycerol). A sample (5 µl) of the digest was separated on a minigel (7 cm x 8 cm) in order to ascertain the level of digestion

achieved prior to the remainder of the reaction (20 µl) mix being separated on a large (20 cm x 25 cm) gel as described below.

Electrophoretic separation of DNA fragments:

RAPD fragments and other amplified DNA products:

RAPD products and other amplified DNA products were separated by electrophoresis on 1.5% MP agarose (Multi-purpose agarose, B. Mannheim) gels made with 1x TBE (90 mM Tris-borate, 2 mM EDTA, pH 8.0) and run at 5v/cm for two hours. A standard DNA size marker (marker VI, B. Mannheim) was loaded at a concentration of 1 µg to provide a relative indication of amplified DNA concentration as well as a means of assessing the size of this amplified DNA. DNA fragments were stained with ethidium bromide, which was added to the gel to a final concentration of 0.5 mg ml^{-1}. The fragments were visualised by applying ultra-violet (uv) light to the gels using a transilluminator (UVP, Inc.).

Genomic DNA and other non-amplified DNA fragments:

Total genomic DNA and the fragments resulting from digestion using restriction enzymes were separated on agarose gels (0.7%) made with 1x TBE. These gels were stained with ethidium bromide and visualised as described above. Restriction digests were run overnight at 2v/cm or until the 4.4 Kb sized band had run about 15 cm. Total genomic DNA was run at 5v/cm for two hours. Size markers were added to both types of gel, the later having 0.5 µg of undigested lambda-phage DNA (from c I 857 Sam 7, B. Mannheim). Size markers used on gels of restriction digests were; (a) one lane containing 0.5 µg of DNA size marker VII (B. Mannheim); (b) one lane of a *Hind*III digest of lambda DNA (Pharmacia, St. Albans, Herts, U.K) (0.5 µg); (c) two lanes of digoxigenin-labelled DNA size Marker II (B. Mannheim) (0.25 µg).

Transfer of DNA to nylon membranes (Southern blotting):

Following electrophoresis, those gels containing DNA for transfer to nylon membranes were depurinated for 10 min in HCl acid (0.25 M) prior to denaturation for 10-15 min in 0.4 M NaOH. DNA was transferred to the nylon membrane (Zeta-probe GT, Biorad, Hemel Hempstead, Herts., U.K.) with the aid of a VacuAid suction blotter (Hybaid, England) at a pressure of 40 cm Hg for 1 hour (as manufacturer's instructions). Transfer of DNA was confirmed by examination of the gel under uv light. The membranes were removed from the VacuAid unit and rinsed briefly in 2x SSC to remove any gel particles before drying on 3 MM filter paper for 30 min. Finally, the DNA was fixed to the membrane by baking at 80° C for two hours.

Construction of Digoxigenin (Dig) labelled probes:

All probes used in this research have been isolated from RAPD profiles thus permitting the use of DNA amplification as a method of labelling.

Isolation of potential probes:

Two methods were used, in this work, to isolate potential species-specific and strain-specific *Rhizobium* probes. The first technique employed was the DEAE paper extraction method described by Sambrook (1989) with the following modifications; DE-81 paper (Whatman International Ltd, Maidstone, Kent, U.K.) was used in place of NA-45 paper and the phenol-chloroform and ammonium acetate extraction steps were omitted. DNA was finally dissolved in 10 µl of 1x TE (10 mM Tris chloride, 1 mM EDTA, pH 7.6). This solution was used to provide a template for DNA amplification.

The second method of probe isolation used was the technique of band-stabbing (Bjourson and Cooper, 1992). This method is described below.

A reaction mixture was prepared as described above but with 25 µl of water in place of 1 µl of cells and 24 µl of water. RAPD profiles were separated on agarose

gels before bands of interest were stabbed vertically through the middle with a sterile 22 gauge needle. The needle was withdrawn from the gel, immediately immersed into the fresh reaction mixture and agitated for 30 seconds by twirling the needle between forefinger and thumb. The needles were discarded and the mixture vortexed for 30 seconds to ensure complete mixing was achieved. The mix was sealed with mineral oil, as for a normal reaction, before DNA was amplified using the profile described above. Following amplification a sample (8 µl) of the amplified band was run on a 1.5% gel to provide a template for the labelling reaction.

Probe labelling:

Probes were labelled with Dig-dUTP (B. Mannheim) using the following procedure. An amplification reaction mixture was made up as described previously but with the exception that the dNTPs (CTP, GTP, ATP) were added to a final concentration of 200 µM each whilst dTTP was added to 182.5 µM and Dig-dUTP to 17.5 µM final concentrations. DNA was added either by the addition of 1 µl of the DNA solution produced following the DEAE paper extraction method or by band-stabbing from the purified band as described above. Labelled probe was produced using the temperature profile described above. Correct labelling was confirmed by running samples (5 µl) of labelled and unlabelled band in adjacent lanes of a minigel and checking for a slower migration of the labelled probe.

Non-radioactive detection of DNA fragments:

Prehybridisation:

Membranes were prehybridised at 68°C in a hybridisation oven (Hybaid, England). Initially membranes were soaked in 2x SSC (1x SSC is 0.15 M NaCl, 0.15 M sodium citrate), prior to prehybridisation for 1 hr at 68° C in prehybridisation solution (5x SSC, 0.1% w/v N-Lauryl sarcosine [Sigma], 1% w/v blocking reagent [B. Mannheim]) (20 ml per 100 cm^2 of filter).

Hybridisation:

The prehybridisation solution was replaced with hybridisation solution (as prehybridisation solution) (2.5 ml per 100 cm^2 of filter). The Dig-labelled DNA probe was denatured by heating at 95° C for 10 min and then added to the hybridisation solution to a final concentration of 20 ng ml^{-1}. Denatured Herring Sperm DNA was also added to give a final concentration of 50 µg ml^{-1}. Hybridisation proceeded at 68° C for 8-12 hrs.

Stringency washing of hybridised filters:

Unbound probe was removed by washing the membranes in a stringency wash buffer. Initial washes were carried out in 2x SSC, 0.1% SDS at 37° C for 2x 30 min. Higher stringency washes, where needed, were carried out at an increased temperature or at a lower salt concentration as described in the relevant chapter.

Detection:

The (3-(2'-Spiroadamantane)-4-methoxy-4-(3"-phosphoryloxy)-phenyl-1,2-dioxetane) (Lumigen-PPD) chemiluminescence detection method (B. Mannheim) was used to visualise probe positioning. Following the stringency washes the membranes were washed briefly (2 min) in Lumigen-PPD Wash Buffer (0.3% Tween 20 [v/v], 0.1 M Maleic acid, 0.15 M NaCl, pH 7.5) before incubation for 30 min in Lumigen-PPD Buffer 2 (1% blocking agent [w/v], 0.1 M Maleic acid, 0.15 M NaCl, pH 7.5) at 100 ml/100 cm^2 of membrane. The membranes were incubated for a further 30 min in Lumigen-PPD Buffer 2 (20 ml/100 cm^2) containing anti-Dig-AP-conjugate (75 mU/ml) (B. Mannheim). The membranes were washed twice for 15 min in 100 ml/100 cm^2 of Lumigen-PPD Wash Buffer before equilibration in Lumigen-PPD Buffer 3 (0.1 M Tris chloride, 0.1 M NaCl, 50 mM MgCl2, pH 9.5) (20 ml/100 cm^2) for 5 min. Lumigen-PPD (B. Mannheim) was diluted in Lumigen-PPD Buffer 3 to a final concentration of 100 µg ml^{-1} and the membranes were incubated for 5 min in 15 ml/100 cm^2 of this substrate solution. After incubation the membranes were dried

briefly on filter paper, sealed in incubation bags and pre-incubated for 15 min at 37° C. Finally the membranes were exposed to light sensitive film (Hyperfilm-ECL, Amersham) for 25 mins which was developed as recommended.

Reprobing:

Membranes were stripped of old probes by rinsing in distilled water for 2 min then washing twice in 0.2 M NaOH, 0.1% SDS for 15 min at 37° C. Finally the membrane was washed in 2x SSC before being rehybridised.

Tables and Figures

Table 2.1: Strains used in this study.

Strain.[1]	Synonym	Geographical Origin of Host	Species.	Reference
R. leguminosarum				
bv *trifolii*				
JJD4[2]		England	*T. repens*	This thesis
3	BB1	England	*T. repens*	
5	Clover F	England	*T. repens*	
35	Orkney 1	Scotland	*T. repens*	
221	TA1	Tasmania	*T. subterraneum*	Humphrey & Vincent (1969)
0404	A'f 12	Sweden		Nutman (1946)
162S7a[3]		U.S.A.	*T. repens*	
162BB1[3]		Zimbabwe	*T. africanum*	
162P17[3]		U.S.A.	*T. pratense*	
162X7a[3]			*T. subterraneum*	
7D5		England		Mytton & Livesey (1983)
RAC 37[4]	TA1	Tasmania	*T. subterraneum*	Humphrey & Vincent (1969)
JJD15[2]		England	*T. repens*	This thesis
46rif[5]		England		Skot L., (pers. comm.)
R1.3		England		Mytton L.R., (pers. comm.)
RAC 41[4]	Clover F	England	*T. repens*	
JJD9[2]		England	*T. repens*	This thesis
JJD17[2]		England	*T. repens*	This thesis
bv *viciae*				
1001	Dutch pea	Holland	*Pisum* spp.	
1004	GW			
1007				
1011		England		
1012	14	Holland	*P. sativum*	Helz *et al.*, (1927)
1013	V1	England	*Vicia sativa*	Nutman (1959)

1014		England		
1015	V3	England	*Vicia* spp.	
1016	S Dutch Pea	Holland	*Pisum* spp.	
1017		Holland		
1018	310	U.S.A.	*P. sativum*	Allen & Baldwin (1931)
1019		U.S.A.		
1020		U.S.A.		
1021		Holland		
1023		Holland		
1024		Holland		
1025		Holland		
1026		Holland		

bv *phaseoli*

3603				
3604				
3605	CC511	U.S.A.	*Phaseolus vulgaris*	Gemmell & Roughley (1975)
3606				
3607		Denmark		
3608	X1	Brazil	*Mimosa caesalpiniaefolina*	
3609	127E12	U.S.A.	*P. limensis*	
3611		U.S.A.		
3613		Cuba		
3614		Cuba		
3615		Cuba		
3617		England		
3618	RP11	England		
3619		England		
3620		England		
3622	CSYR 35	England	*P. vulgaris*	
3624		U.S.A.		Pessanha *et al.*, (1972)
3626		Brazil		Guss & Dobereiner (1972)

R. meliloti

2000		Denmark		
2001	AH2	Denmark		Thornton

				(1931)
2002		Australia		Gibson & Nutman (1960)
2003		Canada		
2004		Canada		
2005		Australia		
2006		Australia		
2007		Australia		
2008				
2009	Manitoba	Canada		
2010		Australia		
2011	SU47	Australia		Gibson (1959)
2012	SU277	Australia	*Medicago truncatula*	Gibson (1959)
2013		Jordan		
2015				
2016		Czechoslovakia		
2017		Czechoslovakia		
B. japonicum				
3403	507	U.S.A.	*Glycine max*	Allen & Baldwin (1931)
3408		Canada		Dobereiner *et al.*, (1972)
3413		Brazil		
3427	311b110	U.S.A.	*Glycine max*	Means *et al.*, (1964)
3437		Japan		
Bradyrhizobium spp.				
3201	Bulduri 1			Kalnins (1938)
3205				
3212		England		Witty & Day (1977)
3213	NZ 4142	N. Zealand	*Lupinus arboreus*	
3801			cowpea	
3819		U.S.A.	*Dolichos* spp.	
3827	CB 1189	U.S.A.	*Cicer arietinum*	Islam &

3828	U.S.A.	*Cicer* spp.	Dart (1973) Islam & Dart (1973)

[1] All strains, unless otherwise stated, come from the Welsh Plant Breeding Station (WPBS) collection, Aberystwyth, Wales.

[2] Isolated from white clover (*T. repens*), permanent pasture, The Royal Agricultural College, UK. England.

[3] Strains supplied by the Nitrigin company, U.S.A.

[4] Sub-cultures of strains, originally from WPBS collection, now maintained at the Royal Agricultural College.

[5] Rifampicin resistant strain.

Table 2.2: Rhizobial growth media.

Media	Components	Final Conc.	Reference
TY Agar			Beringer, (1974)
	Tryptone	0.50%	
	Yeast Extract	0.30%	
	$CaCl_2$	6 mM	
	Agar	1%	
PA Solution			Hirsch *et al.*, (1984)
	Peptone	0.40%	
	$MgSO_4 \cdot 7H_2O$	0.05%	

Phylogenetic Grouping and Identification of *Rhizobium* Isolates on the Basis of Random Amplified Polymorphic DNA Profiles.

Material from this chapter has been included in the paper:

DOOLEY J.J., HARRISON S.P., MYTTON L.R., DYE M., CRESSWELL A., SKØT L. & BEECHING J.R. (1993). Phylogenetic grouping and identification of *Rhizobium* isolates on the basis of random amplified polymorphic DNA profiles. Canadian Journal of Microbiology. **39**, 665-673.

Material from this chapter has also been used in the paper:

DYE M., SKØT L., MYTTON L.R., HARRISON S.P., DOOLEY J.J. & CRESSWELL A. (1995). A study of *Rhizobium leguminosarum* biovar *trifolii* populations from soil extracts using randomly amplified polymorphic DNA profiles. Canadian Journal of Microbiology. **41**, 336-344.

Abstract

Using a single, random 15-mer as a primer, between 1 and 12 DNA amplification products were obtained per strain from a selection of 84 *Rhizobium* and *Bradyrhizobium* isolates. A Principal Coordinate analysis was used to analyse the resulting amplified DNA profiles and it was possible to assign isolates to specific groupings. Within the species *Rhizobium leguminosarum*, the biovar *phaseoli* formed a distinct group apart from the other biovars of the species, *viciae* and *trifolii*, which congregated together. Isolates of *Rhizobium meliloti* and *Bradyrhizobium* species formed their own clear, specific assemblages. Although it was possible to identify individual isolates on the basis of differences in their amplified DNA profiles, there was evidence that some amplified segments were conserved among individuals at the biovar and species levels.

Introduction

Various methods employed previously to classify and identify *Rhizobium* strains have included polyacrylamide gel electrophoresis of isoenzymes (Mytton *et al.*, 1977; Young, 1985; Harrison *et al.*, 1989a) and plasmid profiling (Glynn *et al.*, 1985; Bromfield *et al.*, 1987; Harrison *et al.*, 1989a,b). Studies of DNA homology have led to the development of strain-specific probes which can distinguish among different isolates of *Rhizobium loti* (Bjourson and Cooper, 1988). Restriction Fragment Length Polymorphisms (RFLPs) have been used to study symbiotic DNA from *R. leguminosarum* bv *trifolii* isolates (Watson and Schofield, 1985; Harrison *et al.*, 1988) and specific probes to symbiotic DNA have been further used by Schofield *et al.*, (1987), Young and Wexler (1988) and Engvild *et al.*, (1990) to examine differences in the three biovars of *R. leguminosarum*. Segovia *et al.*, (1991) and Young *et al.*, (1991) have recently examined *Rhizobium* strain relationships using DNA amplification and sequencing of the 16S Ribosomal DNA. These methods have been described more fully in Chapter 1.

The use of random primers to produce Randomly Amplified Polymorphic DNA (RAPD) fingerprints from strains of *R. leguminosarum* bv *trifolii* has been reported by Harrison *et al.*, (1992). In that study amplification profiles were obtained from twelve isolates using a series of random primers. The value of the RAPD fingerprints, from each primer, for strain identification was assessed. From the results of this work the primer, SPH1, was identified as showing the greatest potential for differentiating among the *Rhizobium* strains. This primer was therefore selected for use in this study to examine molecular relationships among strains, biovars and species in a collection of eighty four isolates of *Rhizobium* and *Bradyrhizobium* of varying geographic and phylogenetic origin.

Materials and Methods

Seventy one isolates from *R. leguminosarum* biovars *trifolii*, *viciae* and *phaseoli*, *Rhizobium meliloti* and thirteen strains from the genus *Bradyrhizobium* were employed in this investigation.

Production of RAPD profiles:

RAPD profiles were produced as described in the General Materials and Methods (Chapter 2) using the primer SPH1 (5'-GACGACGACGACGAC-3'), (ILS [International Laboratory Services], Team Valley Trading Estate, Gateshead, Tyne & Wear, NE11 0LH).

Statistical analysis of results:

At the outset of this research, little work was published on the use of RAPDs for phylogenetic studies of any species thus it was necessary to develop a method for data analysis. The use of Principal Components Analysis (PCA) and Principal Coordinates Analysis (PCO) is documented as useful as an initial method of extrapolating underlying trends from data to ascertain its potential use for further analysis. The two methods differ in the type of data to which they are applied, PCA relying on quantitative data and PCO on qualitative data. The very nature of the RAPD reaction produces data of the qualitative type (the bands are either present [1] or absent [0] and cannot easily be quantified from one reaction to the next as an increase in one band intensity leads to a decrease in the intensity seen in other bands) therefore necessitating the application of PCO analysis to the data. Further support for the use of PCO to analyse data composed of '1s' and '0s' is provided by Digby and Kempton (1987) who applied PCO analysis to data from a study of the presence (1) or absence (0) of mosses growing on a shore-line.

RAPD profiles produced using the primer SPH1 were analysed, using PCO and Cluster Analysis, as described in Chapter 2.

Further statistical analysis - the Jaccard matching coefficient :

As a comparison to the simple matching coefficient method of analysis the Jaccard coefficient of matching method was also employed to analyse the RAPD fingerprint data from the 84 strains.

The Jaccard matching coefficient (Jaccard, 1901) varies from the simple matching coefficient in that the Jaccard method eliminates those positions which show the absence (0) of a band in both strains being considered. The simple matching coefficient does not disregard these positions. The Jaccard method therefore returns a similarity of 100% for those positions (loci) which display the presence (1) of a band in both isolates. This method returns a similarity of 0% for those loci which display the presence of a band in one isolate while it is absent from the other isolate. When both isolates lack a band at a particular locus this position is omitted from consideration and plays no part in the calculation of similarity. A comparison of the calculation of similarity (S) using the simple matching (S_{sm}) and the Jaccard matching coefficients (S_J) is shown below:

For two strains i and j at a locus X there are four (a-d) possible band pattern combinations;

a : i = j = 1 i.e both strains have the band present

b : i = j = 0 i.e. both strains show an absence of the band

c : i = 1, j = 0 i.e. strain i has the band while strain j does not

d : i = 0, j = 1 i.e. strain i does not have the band while strain j does.

Similarity (S) is then calculated using the Jaccard coefficient as:

$$S_J = a / (a+b+c+d)$$

Similarity (S) is then calculated using the simple matching coefficient as:

$$S_{sm} = a+b / (a+b+c+d)$$

Results and Discussion

Amplification of DNA from intact cells of each of the eighty four isolates used in this investigation was achieved using a single 15mer primer, SPH1 (Harrison *et al.*, 1992). The primer has an estimated melting temperature of 50° C (Thein and Wallace, 1986) therefore, using an annealing temperature of 36° C should result in some primer/target mismatch. This mismatch is apparently constant, and does not detract from the analysis as the same patterns can be reproduced from individual strains upon subsequent amplifications. It is useful to allow consistent mismatch to occur since the small bacterial genome would not provide sufficient exact target sites to give rise to meaningful profiles if strict annealing temperatures were adhered to.

The resulting amplification profiles for each of the strains are displayed in Figure 3.1, which illustrates those for *R. leguminosarum* biovar *viciae* (a), biovar *trifolii* (b), biovar *phaseoli* (c), *R. meliloti* (d) and *Bradyrhizobium* (e) respectively. Comparison of amplified segments against size markers using electrophoresis allowed estimation of their lengths.

All amplified products were between 0.1 Kb and 3.0 Kb in length and the number of bands per strain varied between 1 and 12. From this variation in number and size of bands it is possible to identify individual strains from what could be called a "DNA fingerprint". Although strains for the most part give rise to unique identifiable profiles, the patterns produced by several strains can be seen to be very similar; Figure 3.1b, *R. leguminosarum* biovar *trifolii* strains *Rt*5 and *Rt*221; Figure 3.1c, *R. leguminosarum* biovar *phaseoli* strains *Rp*3613 and *Rp*3615; Figure 3.1d, *Rhizobium meliloti* strains *Rm*2001, *Rm*2005 and *Rm*2008, and Figure 3.1e, *Bradyrhizobium* strains *Br*3205, *Br*3212 and *Br*3213). The similarity in amplification profiles of different strains suggests a certain degree of DNA conservation among isolates. It is also possible that different strains from different culture collections are identical, and thus give rise to similar amplification products.

The *Bradyrhizobium* strains contain 5 unique bands which are not seen in the *Rhizobium* species. One of these bands of size 0.48 Kb is present in 8 of the 13 strains suggesting DNA conservation within the *Bradyrhizobium* genus. Similar observations of bands common to nearly all strains of the *R. leguminosarum* biovar *trifolii* and biovar *viciae* species can also be made. Moreover, the presence of three different common bands in some *R. meliloti* strains (Figure 3.1d, strains *Rm*2000, *Rm*2001, *Rm*2003, *Rm*2004, *Rm*2005, *Rm*2007, *Rm*2008) suggests strong DNA conservation within this species.

Analysis based on the presence or absence of amplified products:

A statistical analysis of band sizes within and between biovars and species could be useful in assigning strains to species and sub-specific groups based on the occurrence of conserved fragments. This could also provide a means assessing the level of relatedness between these species and biovars.

When all amplified products from all strains were added together forty-five different products were shown to exist in total across all isolates examined. These were converted into a presence (1)/absence (0) file, consisting of a string of 45 '1's or '0's per isolate. This file is displayed in Appendix 2.

A similarity matrix was constructed from the presence/absence data, using both the simple matching and Jaccard matching coefficients, as described in Chapter 2. These matrices are shown in Appendices 3 and 4 respectively.

Application of Principal Coordinates Analysis (PCO):

From a plot of the first two ordination axes (ordinate 2 against ordinate 1), resulting from analysis of the amplification products from the eighty four isolates, it appeared that a series of groups, relating to individual species and biovars, were emerging. Following the application of Cluster Analysis (average linkage) it was possible to denote clusters as areas of coloured shading, according to their percentage relationship (ranging from 70-100%), on the PCO plot. These clusters when applied

confirmed the previous suspicions that groups were emerging and thus indicated that differentiation between the various species is possible based on the presence or absence of amplified bands. The final ordination produced is shown in Figure 3.2. The first two axes shown account for the greatest amount of variation found within the data, which in this case is 24%. That means 24% of the total variation among all isolates is described by the first two axes.

This Figure shows that distinction of the *R. leguminosarum* biovars *viciae* and *trifolii* from both the *R. meliloti* and *Bradyrhizobium* is achieved though separation along axis 1. All *R. meliloti* and *Bradyrhizobium* are located in the region of the graph bounded by the scores -0.4 to 0.0 on axis 1 whereas the majority of the *R. leguminosarum* biovar *viciae* and biovar *trifolii* have scores of greater than 0. *R. leguminosarum* biovar *phaseoli* strains form a fairly tight cluster and are located towards the centre of axis 1, having values of between -0.1 and 0.1. However there is a certain amount of overlap with the other clusters. Consequently biovar *phaseoli* cannot be differentiated using this dimension alone.

Examination of the second axis of the ordination allows for differentiation between the species *R. meliloti* and the genus *Bradyrhizobium*, the former mainly having values greater than -0.02 and the latter less than -0.08. *R. leguminosarum* biovar *phaseoli* isolates can be differentiated by high scores on axis 2; with all but one exception these strains have values exceeding 0.13. Only two non-*R. leguminosarum* biovar *phaseoli* strains are found within that region of the ordination plot that exceeds 0.13 with respect to axis 2. However preliminary investigations suggest that the *R. meliloti* strain may be mislabelled as it failed to nodulate Lucerne (*Medicago sativa* L) plants. Using the PCO plot it is not possible to differentiate between *R. leguminosarum* biovars *viciae* and *trifolii* with any degree of certainty.

From the shading resulting from the Cluster Analysis applied to the isolates described in Figure 3.2, it can be seen that eight clusters are present at the 80-100% level (denoted by a red line on Figure 3.2). This level of similarity indicates that a very high degree of DNA conservation exists among the isolates within these gatherings. Of

these eight collections three account for isolates of *R. leguminosarum* biovar *phaseoli* only, one contains mainly *R. leguminosarum* biovar *phaseoli* isolates with the addition of one *R. meliloti* isolate and one isolate each of the biovars *viciae* and *trifolii*. The remaining four clusters comprise either *R. meliloti* or *Bradyrhizobium* isolates.

The *Bradyrhizobium* strains fall into two clusters although a single isolate (isolated from lupins) associates with neither group at the 80-100% level. The right hand cluster comprises three strains of the recognised species *B. japonicum* (based on a recent review of current Rhizobial classification, Elkan, 1992) and an unclassified *Bradyrhizobium* strain. The left hand group comprises three strains isolated from lupins and two unclassified *Bradyrhizobium* strains. This observation would appear to indicate that within the genus *Bradyrhizobium* there is more than one species detectable at the 80-100% level of similarity. This agrees with results found by other researchers (Stanley *et al.*, 1985; Kuykendall *et al.*, 1988) and summarised in a mini-review by Elkan, (1992). At the 75-80% level of similarity (denoted by a blue line on Figure 3.2) the lupin-derived *Bradyrhizobium* isolate associates with the *B. japonicum* cluster. This suggests that the variation, observed at the 80-100% level of similarity, between the two potential species within the genus *Bradyrhizobium* may not be completely consistent. However thirteen *Bradyrhizobium* isolates may be too small a sample to provide sufficient data to allow an accurate analysis to be carried out.

A division within the species *R. meliloti* at the 80-100% level has resulted in the formation of two groups (excluding the isolate in the *R. leguminosarum* biovar *phaseoli* cluster described earlier) of eleven and four isolates respectively. This division of the species into two clusters may indicate that either the *R. meliloti* species comprises two distinct sub-species or is in fact two separate species. Eardly *et al.*, (1990) have reported similar indications based on observations of electrophoretic variation in the alleles of 14 metabolic enzymes. They confirmed this with data from restriction fragment length polymorphisms (RFLPs) and concluded that *R. meliloti* should be regarded as two separate species. Young (1985) has also reported the existence of two major electrophoretic types (ETs) in *R. meliloti* reinforcing this idea.

Within the cluster comprising *R. leguminosarum* biovar *phaseoli* (as defined at the 70-75% level) there are four highly related groups (80-100% similarity), three of which include only two isolates while the fourth encompasses 10 strains. This observation suggests that although this biovar appears to be formed from a single species it may in fact be an amalgamation of several species, or sub-species, which are only detectable at higher relationship levels. These observations are similar to those of Pinero *et al.*, (1988) who have suggested that *R. leguminosarum* biovar *phaseoli* is a polyphyletic group comprising at least seven individual species. It is possible that the work in this study has resulted in the identification of four sub-species groupings within *R. leguminosarum* biovar *phaseoli*. Even from simple, subjective observations of band patterns (Figure 3.1c) it would appear that the strains within *R. leguminosarum* biovar *phaseoli* are not as genetically related to each other as those strains comprising the biovars *R. leguminosarum* biovar *trifolii* and biovar *viciae*. These differences have recently led to the formation of a novel species, *R. tropici* (Martinez-Romero *et al.*, 1991), which encompasses all those strains formerly known as *R. leguminosarum* biovar *phaseoli* Type II.

The additional information obtained on the *R. leguminosarum* biovar *phaseoli* strains used for this study is not sufficient to allow for reclassification of the strains into the species *R. leguminosarum* biovar *phaseoli* or *R. tropici*.

At the 75-80% similarity level several new clusters are evident, two of which are composed of a mixture of *R. leguminosarum* biovar *trifolii* and biovar *viciae* isolates. These two clusters account for most of the strains of these two biovars. Within the *R. leguminosarum* biovar *phaseoli* isolates a small assembly binds two of the more highly related groups, as defined at the 80-100% level, thus indicating a closer relationship exists between these two clusters than exists between the other clusters and isolates comprising the *R. leguminosarum* biovar *phaseoli* species.

Of the remaining two groups, formed at the 75-80% level, one links a single *Bradyrhizobium* isolate to the *B. japonicum* cluster whilst the other links the left hand group of *Bradyrhizobium* strains with one of the *R. meliloti* clusters thereby

suggesting a closer relationship exists between these two, seemingly unrelated, species than exists between the two sub-groups of the species of *R. meliloti*. This observed difference may be the result of a small variation being exaggerated owing to a low number of strains being examined.

At the lowest level of clustering considered in this investigation, the 70-75% level, (denoted as a black line on Figure 3.2) four major clusters are seen to result. These consist of *R. leguminosarum* biovars *trifolii* and *viciae*, *R. leguminosarum* biovar *phaseoli*, *R. meliloti* and *Bradyrhizobium*. There is however a small amount of overlap between the clusters defining the biovars of the *R. leguminosarum* species so suggesting that although spatially separated in Figure 3.2, there is still a certain amount of conserved DNA linking the isolates of this species.

Only one group occurs at the 60% level encompassing all isolates. However at the 65% level, two clusters exist one being the *B. japonicum* group and the other a large group comprising the other isolates.

From Figure 3.2 the following conclusions can be drawn: (i) Differences occur at the molecular level among the species *R. leguminosarum*, *R. meliloti* and the genus *Bradyrhizobium*. Demezas *et al.*, (1991) have previously reported differences between *R. leguminosarum* biovar *trifolii* and *R. meliloti* illustrated by studies of allozymes and RFLPs; (ii) There is no absolute difference between *R. leguminosarum* biovar *trifolii* and *R. leguminosarum* biovar *viciae* that can be found using this method of analysis with this sample of strains; (iii) A difference between the biovar *phaseoli* and the other biovars, *trifolii* and *viciae*, of the species *R. leguminosarum* does occur although variation is reduced at the 70-75% level as a certain amount of overlapping of clusters occurs suggesting some similarity among these biovars. From their work on electrophoretic types in non-symbiotic isolates, Segovia *et al.*, (1991) have reported similar observations, with differences occurring between *R. leguminosarum* biovar *phaseoli* type I strains and *R. leguminosarum* biovar *trifolii*, *viciae* and *R. tropici* strains. Martinez-Romero and Rosenblueth (1990) have also reported differences between host range, *nif* gene reiterations and competitiveness in strains of *R.*

leguminosarum biovar *phaseoli* (type I) and *R. tropici*. Reports by Johnston and Beringer (1977) and Kondorosi *et al.*, (1980) of readily obtainable chromosomal genetic exchange among the biovars *trifolii*, *viciae* and *phaseoli*, of the species *R. leguminosarum* could account for their greater relationship to each other than to *R. meliloti* which would appear not to readily exchange chromosomal material with *R. leguminosarum* strains. (iv) Although distinct relationships have been found, a few strains did not fall within their expected groupings. For instance one *R. meliloti* strain occurred in the *R. leguminosarum* biovar *phaseoli* cluster and one *R. leguminosarum* strain occurred in the *R. meliloti* cluster. There is also close affinity between a collection of *Bradyrhizobium* strains and one of the *R. meliloti* groups. These observations agree with those of Young (1985) and Young *et al.*, (1987) who suggest that genetic boundaries among nodulating *Rhizobium* strains can be somewhat indistinct. Our results indicate that this may also apply to *Bradyrhizobium* although this would need confirming by examining a larger and more diverse sample of strains.

This method of data analysis and the results produced have been published, Dooley *et al.*, (1993).

Jaccard matching coefficient:

The criterion used by this analytical method to produce similarity matrices differs from simple matching coefficient in that the absence of a band at a particular loci in both strains results in this loci being omitted from the analysis. This has been summarised above. This method was used as a comparison to the simple matching coefficient in order to test if the absence of a band, at a particular loci in both strains, is as important as the presence of the band in both strains. The similarity matrix produced using the Jaccard matching coefficient is shown in Appendix 4.

The plot of ordinate 2 against ordinate 1, following the application of PCO to the similarity matrix constructed from analysis using the Jaccard matching coefficient, is shown in Figure 3.3. From an initial observation of the plots (Figures 3.2 and 3.3) it

can be seen that they appear very similar, with the exception that they are mirrored images of one another about, roughly, the origin of ordinate 1. Upon closer examination several smaller differences start to emerge as described below.

The *Bradyrhizobium* isolates (with the exception of one isolate) fall within the lower right-hand quarter of the graph, i.e. that sector where values are greater than 0 with respect to ordinate 1 and less than 0 with respect to ordinate 2. The *R. meliloti* strains all show a similar distribution with the exception of two isolates, the one discussed earlier and shown not to nodulate Lucerne, and a second isolate which lies just above the origin with respect to ordinate 2. All strains however fall on the positive side of the axis for ordinate 1. It can thus be assumed that any strain analysed by this technique and falling within this section of the graph will have a high probability of being either a *Bradyrhizobium* or a *R. meliloti* isolate.

Further observations reveal that all the *R. leguminosarum* biovar *phaseoli* isolates are found in the region of the graph above the origin with respect to ordinate 2. This is an identical observation to that made by analysis with the simple matching method. They also all fall within a narrow band bounding the origin of ordinate 1, all isolates being contained within the region -0.25 to +0.25 units. This again reflects results obtained from simple matching.

The *R. leguminosarum* biovar *viciae* and *R. leguminosarum* biovar *trifolii* isolates can be seen to occupy, on the whole, the lower left-hand section of the plot, Figure 3.3. Again, as with the analysis using simple matching, it is not possible to distinguish between isolates of *R. leguminosarum* biovars *viciae* and *trifolii*. Most strains have values of less than 0.1 with respect to both ordinate 1 and ordinate 2. The only exceptions to this being two *R. leguminosarum* biovar *viciae* isolates and one *R. leguminosarum* biovar *trifolii* isolate, all three of which appear to associate with the *R. leguminosarum* biovar *phaseoli* isolates.

From simple observations of the ordinate plots produced using either the Jaccard method of matching (Figure 3.3) or simple matching (Figure 3.2) no overall differences can be seen except for a mirroring of the image. This would indicate that it

does not matter which matching technique is employed for the analysis of RAPD data. However when the ordinate plots (Figures 3.2 and 3.3) are examined, following the application of Cluster Analysis to the PCO plots, a noticeable difference in the formation of groups can be observed between the two plots. In general the number of clusters and level at which clustering occurs is lower when Jaccard analysis is applied as compared to results obtained with simple matching.

At the highest level of clustering, 80-100%, only four groups have emerged with Jaccard analysis. These comprise three *Bradyrhizobium* isolates in one group, three *R. meliloti* in another, two *R. leguminosarum* biovar *trifolii* strains comprise the third cluster and two *R. leguminosarum* biovar *phaseoli* the final group. This compares to the eight clusters seen with simple matching, where even at this high level of relatedness the *R. meliloti* and *Bradyrhizobium* species have been defined and shown to comprise of two sub-groups each, Figure 3.2.

The second level of similarity considered was the 75-80% level. The only change seen at this level of similarity, in Figure 3.3, is the addition of one *R. meliloti* strain to the cluster formed at the previous level of similarity. No other strains link to each other or to any of the groups formed previously and the number of clusters remains at four. The number of isolates linked at this level of similarity is very low when compared to the number which cluster together following simple matching. When simple matching was used the majority of the *R. leguminosarum* biovar *trifolii* and *R. leguminosarum* biovar *viciae* isolates had formed two groups at this level of similarity.

The lowest level of similarity examined with the simple matching data was the 70-75% level. At this level the representative species and biovars were all individually clustered with the exception of the *R. meliloti* and *Bradyrhizobium* isolates which had formed a single group. Data from the Jaccard method of matching, when subjected to Cluster Analysis produces only seven clusters at the 70-75% level of similarity. These comprise the four groups observed at the 75-80% level of similarity and three newly formed clusters of two strains each. Two *R. meliloti* isolates form the first group

whilst two *R. leguminosarum* biovar *viciae* strains form the second. The final cluster is formed by the linking of one strain of *R. leguminosarum* biovar *trifolii* and one isolate of *R. leguminosarum* biovar *viciae*. These results alone suggest that there is little DNA homology between isolates of the same species.

A final Cluster Analysis was applied to the Jaccard data at the lower 50-70% level of similarity. At this level of similarity there are still only seventeen groups formed. Three groups are each formed by the joining of two isolates of *R. leguminosarum* biovar *phaseoli*. The remaining *R. leguminosarum* biovar *phaseoli* isolates are still not clustered. This indicates a poor level of similarity within this biovar, however this can, and has been accounted for by using simple matching. Of the remaining associations three are formed by the grouping of strains of *R. meliloti*. These clusters comprise six, two and two isolates respectively. This clustering suggests that there maybe at least three sub-groups within the species *R. meliloti*, however, results from simple matching analysis and from work by Eardly *et al.*, (1990) have only indicated two sub-groups within this species. Three further clusters of five, two and two strains respectively are formed by isolates of *Bradyrhizobium*. These results again differ from those obtained with simple matching. The remaining clusters formed at the 50-70% level of similarity are composed of *R. leguminosarum* biovar *trifolii* and *R. leguminosarum* biovar *viciae* strains. There are two groups of two *R. leguminosarum* biovar *trifolii* strains, one cluster of two *R. leguminosarum* biovar *trifolii* strains and a single *R. leguminosarum* biovar *viciae* isolate, four clusters of two *R. leguminosarum* biovar *viciae* isolates only and a single group of three *R. leguminosarum* biovar *viciae* isolates. These results allow a certain degree of identification of isolates from the two biovars which is not possible with simple matching. However as the level of similarity is reduced to below 50% the number of groups containing a mixture of *R. leguminosarum* biovar *trifolii* and *R. leguminosarum* biovar *viciae* isolates is increased.

The following conclusions can be drawn from observations of results obtained using the Jaccard method of analysis. (i) A spatial differentiation occurs between the

species *R. leguminosarum*, *R. meliloti* and the genus *Bradyrhizobium*. There is also a difference between the biovar *phaseoli* and the other two biovars (*trifolii* and *viciae*) of the species *R. leguminosarum*. (ii) It is not possible to differentiate between isolates of *R. leguminosarum* biovars *trifolii* and *viciae* based on spatial arrangements. This situation is not clarified even after the application of Cluster Analysis. (iii) The overall level of similarity observed between strains of the same species or biovar is low with this method of analysis. (iv) Group sizes are also small and erratically formed which indicates that either the strains are not related or the statistical method employed is not appropriate for this type of analysis. The occurrence of a high number of ungrouped isolates, even at the 50-70% level of similarity, and previous results reported using simple matching tend to suggest that the problem lies with the statistics.

***Conclusions*:**

From the results described above it can be suggested that the analysis of RAPD profiles by the use of the simple matching coefficient followed by Cluster Analysis is a better method than analysis using the Jaccard matching coefficient followed by Cluster Analysis. The results shown here indicate that analysis using the simple matching coefficient produces results which are as equally as acceptable, if not superior, to those produced by use of the Jaccard matching coefficient owing to the higher levels of similarity observed.

Further reasons for the elimination of the Jaccard method were that it was felt that the absence of a band was just as significant as the presence of a band, at a particular loci, when analysing RAPD data as compared to RFLP data. The presence or absence of bands, although appearing to be simply a matter of the presence (or absence) of a DNA fragment in the genome is in fact controlled by two independent areas or 'loci', i.e. the primer binding sites. A change in either of these two sites can result in the loss of the band from a fingerprint and thus a '0' being recorded for that band size in a particular strain. If we then consider two strains we are in fact considering the environments of four 'loci' (primer binding sites) for each band that the

strains share. The Jaccard method of matching fails to appreciate this and so was eliminated for this reason. It was also felt that the large number of band positions (45) and hence 90 "binding loci" and the large number of strains (84) would compensate for any inaccuracies observed with simple matching.

Following the application of Cluster Analysis to the plots the overall levels of similarity observed between strains with the Jaccard method of analysis are greatly reduced as compared to those seen with the simple matching method of analysis. This is easily observed with the Cluster Analysis shading applied to the ordinate plots, Figures 3.2 and 3.3. The point at which the final merging of clusters occurs is at a level of similarity of 62.1 %, when the simple matching method is employed, but only at 0.8 % similarity with the Jaccard method of matching. A higher level of similarity between strains indicates a greater relationship between them. A similarity of 62.1 % suggests that the strains being considered are fairly related to each other. The fact that all the strains are from a soil environment, form associations with leguminous plants and have the ability to fix N_2 would indicate that they are related to a certain degree. A level of similarity of only 0.8 %, as deduced by the Jaccard matching, indicates the strains are totally unrelated to each other. This is not the case and so lends further support to the argument for using simple matching.

The final piece of evidence to support the use of simple matching is found when the 70 % level of similarity is examined. Results from simple matching show that most strains have been linked to their respective species clusters at the 70% level. However with the Jaccard method only eight clusters have started to develop at this level of similarity. This accounts for only 19 strains which is less than a quarter of those used in this study.

This method of strain classification examines homologies at the molecular level. It is likely that primer-target sequences are spread throughout the bacterial genome and that this gives a representative overview of the total genomic homology which allows a reliable means of distinguishing among isolates at the individual, biovar and species levels. Although useful in itself for the study of populations, observations of

product commonality and polymorphism's arising from *Rhizobium* RAPDs offer the opportunity to develop strain, biovar and species-specific DNA probes based on these amplified segments for further population studies.

Tables and Figures

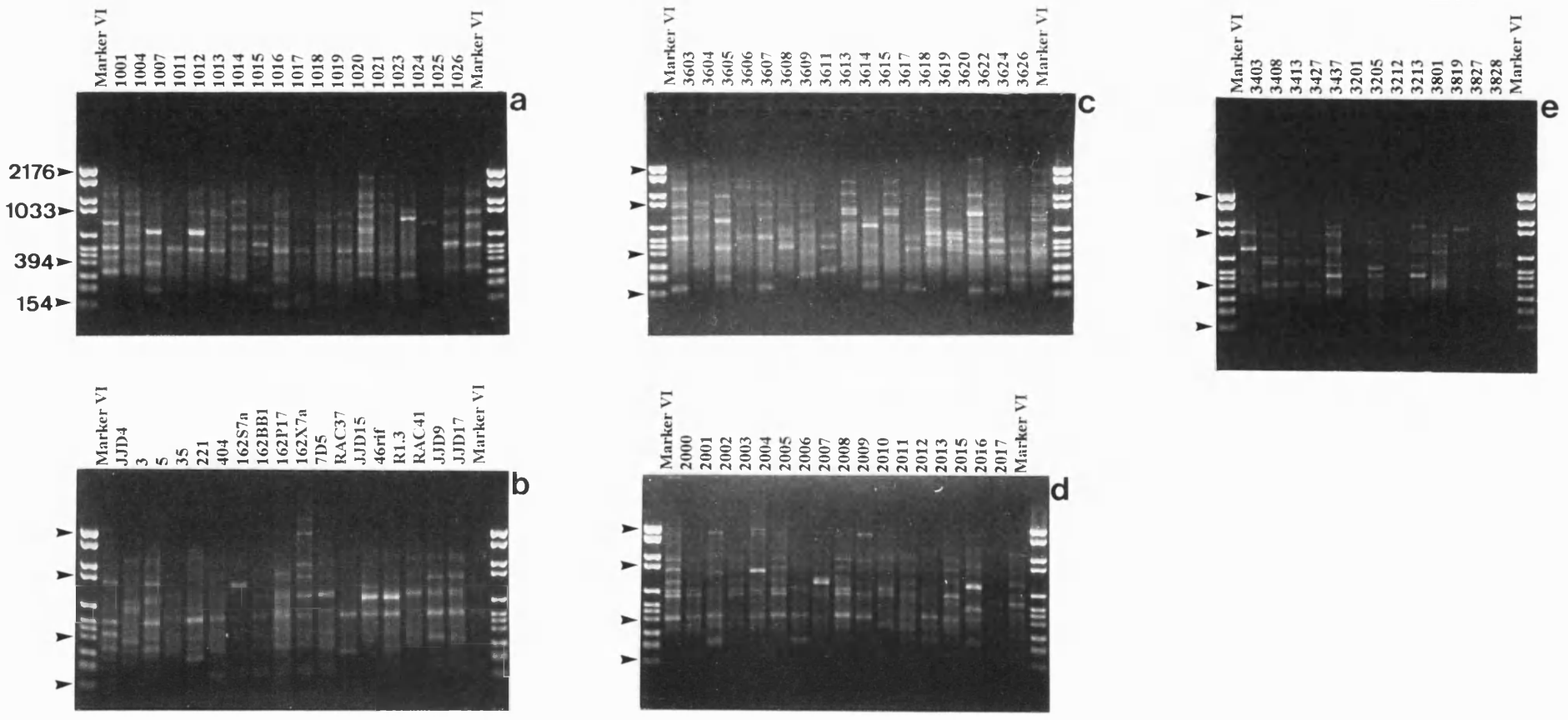


Figure 3.1: RAPD fingerprints produced using the primer SPH1. Patterns were obtained from strains of *R. leguminosarum* biovar *viciae* (a), biovar *trifolii* (b), biovar *phaseoli* (c), *R. meliloti* (d) and *Bradyrhizobium* (e). Strain names are marked above each lane. Marker VI (B. Mannheim) was added to the outer lanes. Marker sizes (arrowed) are in bp.

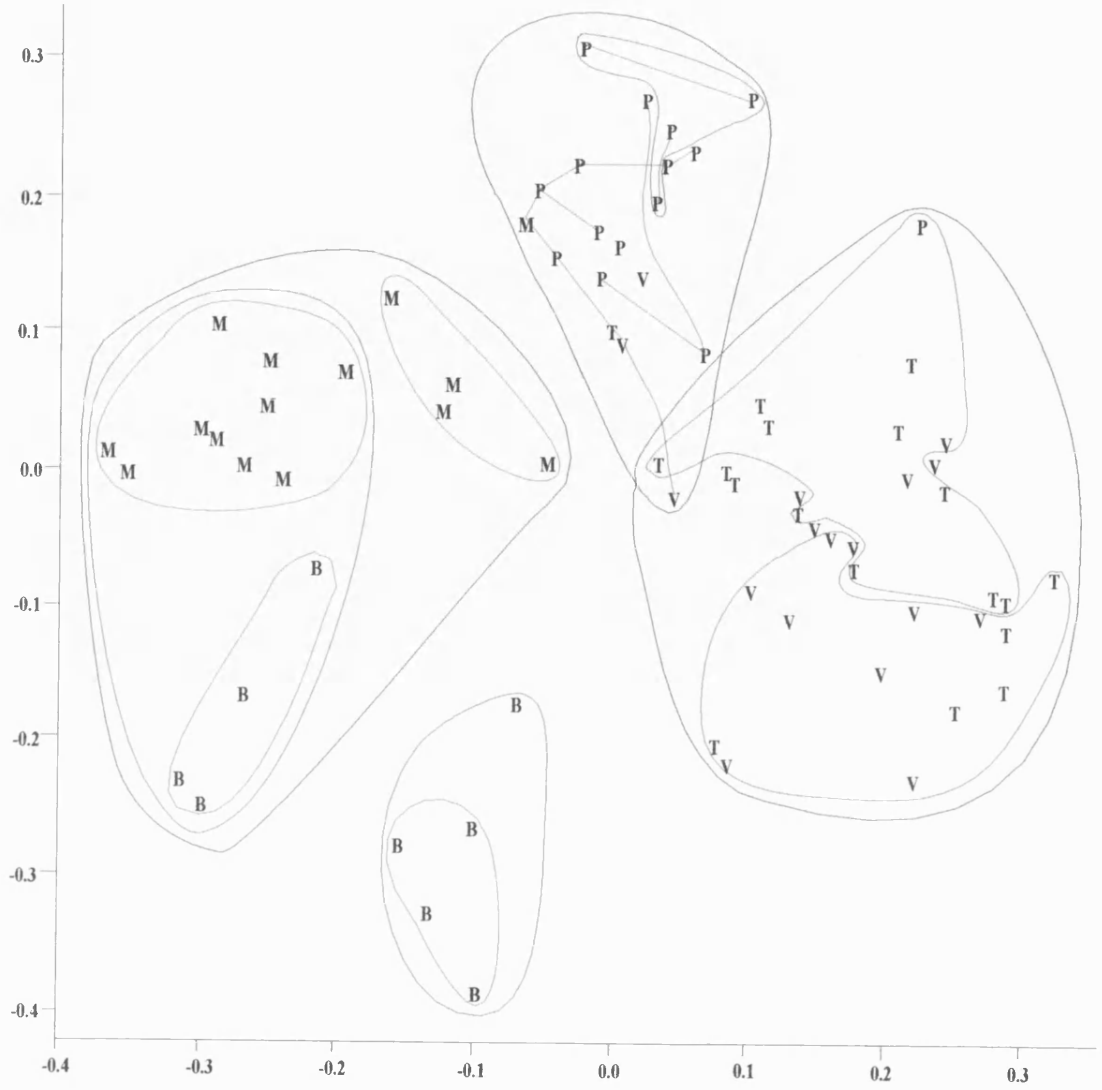


Figure 3.2: Principal Coordinate plot of ordinates 1 and 2 produced following analysis of data from DNA amplification with primer SPH1. Data was matched using Simple Matching. Isolates can be identified as *R. leguminosarum* bv *trifolii* (T), bv *viciae* (V), bv *phaseoli* (P), *R. meliloti* (M) and *Bradyrhizobium* (B). Isolates were grouped according to the results of Cluster Analysis which has been applied to the PCO plot using three levels of clustering; 70-75 % (——), 75-80 % (——) and 80-100 % (——).

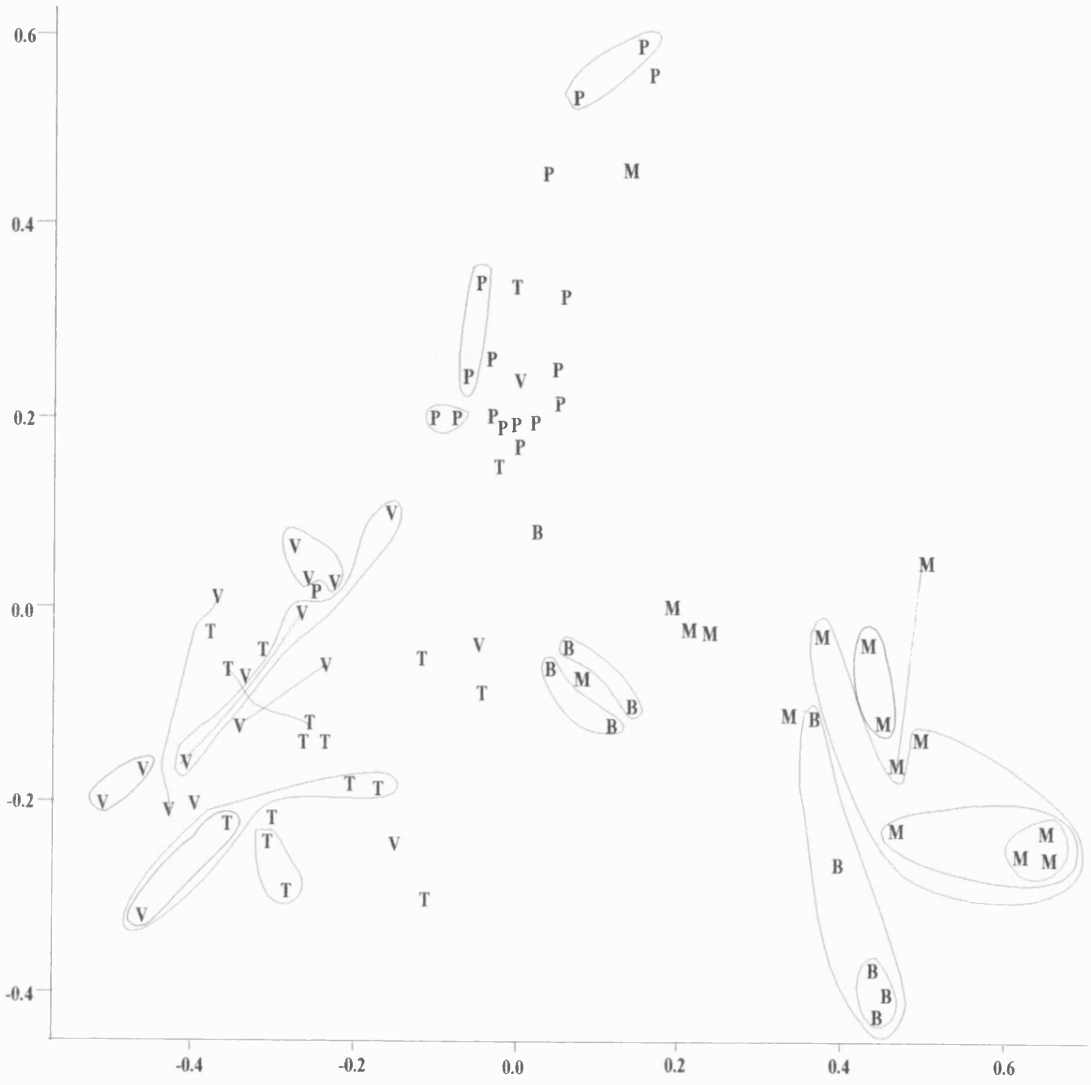


Figure 3.3: Principal Coordinate plot of ordinates 1 and 2 produced following analysis of data from DNA amplification with primer SPH1. Data was matched using Jaccard Matching. Isolates can be identified as *R. leguminosarum* bv *trifolii* (T), bv *viciae* (V), bv *phaseoli* (P), *R. meliloti* (M) and *Bradyrhizobium* (B). Isolates were grouped according to the results of Cluster Analysis which has been applied to the PCO plot using four levels of clustering; 50-70% (——), 70-75 % (——), 75-80 % (——) and 80-100 % (——).

An Investigation of *Rhizobium* Phylogenetics Based on RAPD Profiles Obtained Using Several Different Random Primers

Abstract

From a statistical analysis of RAPD patterns produced using two random primers (SPH3 and SPH7) alone or in dual primed reactions it was not possible to totally differentiate amongst the eighty-four *Rhizobium* and *Bradyrhizobium* isolates under investigation. However from an analysis of patterns produced when using the two primers in a dual primed reaction it was possible to distinguish between isolates of the species *R. leguminosarum* biovar *trifolii* more clearly than when any other primer (including SPH1) was used alone. This suggests that, with the use of the correct primers in a dual primed reaction, it should be possible to classify any *Rhizobium* isolate. When data from several reactions were combined no additional advantage was gained in classifying the isolates, and even in some cases the combinations of data resulted in an overall reduction in resolving power.

Introduction

The use of two primers in a RAPD reaction has been reported by Caetano-Anollés *et al.*, (1991), Fekete *et al.*, (1992) and Welsh and McClelland, (1991). Using dual primers in producing RAPDs results in a different banding pattern to that observed when the two primers are used alone. This double primer reaction may even be used to produce RAPD profiles from primers which, when used alone, showed no, or limited, band production. Possible reasons for these observations are that two target sites for a single primer are too spaced within the genome to permit DNA amplification to proceed to any degree of success. However, when used together the two primers complement each other with their respective primer binding sites over-lapping. This allows amplification to proceed and results in a product which has two different primer sequences at its termini. Welsh and McClelland (1991) and Fekete *et al.*, (1992) report finding that 67% and 30-50% respectively of bands in a double primed reaction are not observed in reactions when each primer is used alone. This therefore indicates that between 30-70% of bands in a double primed reaction have different primer sequences at their termini.

For further analysis of the *Rhizobium* genome it was decided to investigate the efficiency of using two primers (in a double primed reaction) to produce RAPD profiles. The possibility of using these fingerprints for the purposes of classification was also assessed. The selection of the two primers, SPH3 and SPH7, for use in double primed and single primed reactions was based upon the results of a previous investigation by Harrison *et al.*, (1992). This group screened 21 various random primers for their ability to produce amplified products from *Rhizobium* isolates and found that only seven of the primers were capable of producing patterns. Of these seven primers SPH1 was found to be the most efficient whilst SPH3 and SPH7 were slightly less so. These two primers were therefore deemed to offer the greatest potential for producing amplified banding patterns which were not going to be excessively affected by either of the primers.

Additionally, data from two or three amplification reactions using different primers were combined. This was done in order to examine the effect that each data sets may have on the overall results. It was hoped that the use of two or more sets of data would produce a PCO plot which was the result of extrapolation of the strongest trends observed within each individual data set. This, it was hoped, would give a clearer picture of the phylogenies within the *Rhizobium* species.

Materials and Methods

The 84 strains selected for analysis with the primer SPH1 were again employed for analysis with the two primers, SPH3 and SPH7. These strains are shown in Table 2.1, Chapter 3.

Production of RAPD fingerprints:

Amplification reactions were carried out as described previously in the general Materials and Methods (Chapter 2). The primers SPH3 (5'-GACGACAGCGGC-3') (ILS, Gateshead) and SPH7 (5'-CAGCCACAGCGC-3') (ILS), whether used in single primer or double primer reactions, were added to final concentrations of 27 and 29 pmol respectively.

Fingerprints obtained by amplification using these primers were analysed using the techniques described in Chapter 2.

Further analysis:

Further analysis of the data was carried out by combining the data relating to band presence (1) or absence (0), obtained from RAPD profiles produced using the three primers SPH1, SPH3 and SPH3+7 individually. Data were combined as follows; SPH1 and SPH3; SPH1 and SPH3+7; SPH3 and SPH3+7 and finally SPH1, SPH3 and SPH3+7. It was noted that in some combinations, namely those involving SPH3 and SPH3+7, that some bands may be counted twice owing to their amplification in both the single primer (SPH3) reaction and the double primer (SPH3+7) reaction. It was felt that the small number of bands of this type would not unduly effect the results obtained due to the large, overall number of bands being considered. PCO and Cluster Analysis were then performed, as described in Chapter 2, on the resultant new larger data set.

Results and Discussion

Analysis of SPH3 primed reactions:

The results of amplification with the primer SPH3 alone are shown in Figure 4.1 which shows RAPD patterns from isolates of *R. leguminosarum* biovar *viciae* (a), *R. leguminosarum* biovar *trifolii* (b), *R. leguminosarum* biovar *phaseoli* (c), *R. meliloti* (d) and *Bradyrhizobium* (e).

DNA amplification with SPH3 produced between 1 and 20 amplified products per strain with the exception of *R. leguminosarum* biovar *phaseoli* strain 3614 which produced no bands. Band sizes were between 0.3 Kb and 3 Kb. The total number of discernible bands examined between the 84 strains was 60 (see Appendix 5). Analysis of these bands resulted in the production of the similarity matrix shown in Appendix 6.

Figure 4.2 shows the results of applying PCO analysis and Cluster Analysis to the data derived using the primer SPH3 alone. The spatial distribution of isolates within the ordinate plot offers little indication in identifying unknown isolates except on a very general scale. There are few clear groupings of strains from particular species except for two large, ill-defined areas; the lower left-hand quadrant of the graph (-0.4 to -0.1 and -0.3 to 0.2 with respect to ordinates 1 and 2 respectively) and the area of the graph to the right of -0.1 with respect to ordinate 1. The lower left-hand area contains a large proportion of *R. meliloti* and *R. leguminosarum* biovar *phaseoli* strains, with only four other isolates present (three isolates of *R. leguminosarum* biovar *trifolii* and one isolate of *R. leguminosarum* biovar *viciae*). On application of Cluster Analysis, two of the *R. leguminosarum* biovar *trifolii* strains cluster with other groups leaving the remaining *R. leguminosarum* biovar *trifolii* isolate linked with a *R. meliloti* strain. The *R. leguminosarum* biovar *viciae* isolate clusters with a group composed primarily of *R. leguminosarum* biovar *phaseoli* isolates. The second area (to the right of -0.1 with respect to ordinate 1) contains the majority of the isolates from *R. leguminosarum* biovars *trifolii* and *viciae*. However, it

also contains a selection of strains from the other species and biovars thus making absolute identification by simple spatial positioning impossible.

Following the application of Cluster Analysis to the PCO plot (depicted as shading on Figure 4.2) a complex pattern of related isolates develops especially in the region of the plot where there exists a high concentration of *R. leguminosarum* biovar *viciae* and *trifolii* isolates.

When the level of similarity used in the Cluster Analysis is 80-100%, twenty small clusters are apparent, the largest comprising nine strains of mixed species (four isolates of *R. leguminosarum* biovar *trifolii*, one *R. leguminosarum* biovar *viciae* isolate, three *R. meliloti* isolates and one *Bradyrhizobium* isolate). Many of the remaining clusters are composed of two or three isolates of the same species, although seven groups are formed by the clustering of isolates of mixed species.

At the next level of clustering considered (75-80%) only nine clusters are in evidence; one large group containing the majority of strains and eight smaller clusters. The larger group brings together many of the smaller groups formed at the 80-100% level. Several isolates which had remained unlinked were also incorporated into this group. This large cluster contains isolates from all the species and biovars under examination. Of the remaining groups, one links three *Bradyrhizobium* isolates, at the top of the graph, to a single *R. leguminosarum* biovar *trifolii* isolate, another contains six isolates of *R. meliloti* and one *R. leguminosarum* biovar *phaseoli* isolate and there are two groups of three *R. leguminosarum* biovar *phaseoli* isolates only. Two other clusters of three isolates link two *Bradyrhizobium* to a *R. meliloti* and a *R. meliloti*, *R. leguminosarum* biovar *phaseoli* and *R. leguminosarum* biovar *viciae* together. One cluster links two previously free strains, a *R. leguminosarum* biovar *trifolii* and a *R. meliloti*, while the final grouping links a single *R. leguminosarum* biovar *phaseoli* isolate to the largest cluster (formed at the 80-100% level of similarity) forming a cluster of ten isolates. Many of the strains in this group are spatially interspersed within the large group formed at the 75-80% level of clustering.

The final level of clustering considered was the 70-75% level at which only six groups, two formed at the 75-80% level of clustering and four new groups, remain. A single, unlinked *Bradyrhizobium* strain is also present at the 70-75% level of clustering. The two existing groups contain three *Bradyrhizobium* isolates and one *R. leguminosarum* biovar *trifolii* isolate; and an isolate of *R. meliloti* linked to a *R. leguminosarum* biovar *trifolii* isolate, respectively. The third cluster formed at the 70-75% level of similarity is created by the linking of the ten isolate cluster, described at the 75-80% level, to the large group which was also formed at the 75-80% level of similarity. Another cluster is formed when one of the groups of three *R. leguminosarum* biovar *phaseoli* isolates links with a lone *R. leguminosarum* biovar *phaseoli* isolate and the group containing single strains of *R. meliloti* and *R. leguminosarum* biovars *phaseoli* and *viciae*. A ten strain cluster is formed by the linking of two groups, one containing six isolates of *R. meliloti* and one *R. leguminosarum* biovar *phaseoli* isolate and the second consisting of three *R. leguminosarum* biovar *phaseoli* isolates only. The last group formed at the 70-75% level links a lone *R. meliloti* to the group containing two *Bradyrhizobium* and a *R. meliloti* to form a four isolate cluster.

From the statistical analysis of the RAPDs produced by amplification using the primer SPH3 alone it is not possible to identify or classify the various biovars or species with any degree of certainty. There is only a limited amount of separation observable within the PCO plot, this mainly being the *R. meliloti* and *R. leguminosarum* biovar *phaseoli* isolates which have separated from the majority of the other strains. This kind of separation is better observed using the primer SPH1 alone (see Chapter 3). It can be concluded that amplification with the primer SPH3 alone produces no useful data for the classification of the *Rhizobium* and *Bradyrhizobium* species.

Analysis of SPH7 primed reactions:

Amplification with the primer SPH7 alone produced the results shown in Figure 4.3. Profiles were obtained from the following species *R. leguminosarum* biovar *viciae* (a), *R. leguminosarum* biovar *trifolii* (*b*), *R. leguminosarum* biovar *phaseoli* (c), *R. meliloti* (d) and *Bradyrhizobium* (e). It was noted that only ten isolates of *Bradyrhizobium* gave rise to RAPD profiles using this primer alone. No pattern was obtained from the isolates *Br*3201, *Br*3427 or *Br*3819. As can be seen, from Figure 4.3, a limited number of bands per strain (between 1 and 9 [isolate *Rp*3609]) were produced, of which at least one displayed a high level of homology throughout the species. This limited number of bands offered no potential for classification of the species therefore no further analysis using this data was attempted. It was noted, however, that a common band of 660 bp size may be of use as a probe for the detection of strains of either *Rhizobium* or *Bradyrhizobium*. A second band of size 977 bp was observed in the majority of strains from the species *R. leguminosarum*, including those of the biovar *phaseoli*. This band may be of use for detection of strains from this species. Neither lines of enquiry regarding the potential probes were followed further in this study.

Analysis of SPH3+7 primed reactions:

Fingerprints produced by amplification using the two primers SPH3+7 are shown in Figure 4.4 which shows patterns from *R. leguminosarum* biovar *viciae* (a), *R. leguminosarum* biovar *trifolii* (b), *R. leguminosarum* biovar *phaseoli* (c), *R. meliloti* (d) and *Bradyrhizobium* (e).

DNA amplification with SPH3+7 produced between two and twenty amplified products per strain. The product sizes observed with the double primed reaction were between 0.09 Kb and 3 Kb. The average product size from the double primed reaction was generally smaller than that observed in reactions where the primers were used singly. Similar results were reported by Caetano-Anollés *et al.*, (1991), Fekete *et al.*, (1992) and Welsh and McClelland, (1991). The total number of bands shared among

all 84 strains was 85 (see Appendix 7). This number is greater than that observed with either of the primers alone and is probably accountable by the increase in the number of smaller products produced in this double primed reaction. Of these 85 bands only 44 (52%) are similar to those found when each primer is used alone. This observation is similar to that of Fekete *et al.*, (1992) who found between 50-70% of amplified products in a double primed reaction are common to reactions when either primer is used alone. Welsh and McClelland, (1991), however, only found 30% of bands are common to both type of reaction. The similarity matrix produced following analysis of the bands is shown in Appendix 8.

Figure 4.5 shows the ordination plot resulting from analysis of the RAPDs produced from a double primed reaction. From initial observations it can be seen that some spatial separation occurs, in that all *R. leguminosarum* biovar *trifolii* and *R. leguminosarum* biovar *viciae* (except one *R. leguminosarum* biovar *viciae* isolate) strains fall on the side of greater than -0.1 and 0.0 respectively, with respect to ordinate 1. There is also a further separation in that the *R. leguminosarum* biovar *viciae* isolates mostly fall below 0.0 with respect to ordinate 2 whereas the *R. leguminosarum* biovar *trifolii* isolates are mainly located above this intersection. This would suggest that with larger numbers of isolates, of these two biovars, it may be possible to employ this double primed reaction as a tool to identify differences between the two biovars of *R. leguminosarum* which are undetectable using the primer SPH1 alone (see Chapter 3). *R. leguminosarum* biovar *phaseoli* isolates appear to be gathered in the bottom left-hand section of the plot with values of less than 0.0 and -0.05 with respect to ordinates 1 and 2 respectively. Both the species *R. meliloti* and *Bradyrhizobium* have separated into the upper left portion of the plot, with most isolates having values of greater than -0.05 with respect to ordinate 2 and less than 0.05 with respect to ordinate 1. These two species appear to be very intermingled using this plot although a few small clusters are observable.

Upon application of Cluster Analysis the apparent clusters do not hold so well. At the 80-100% level of similarity, one large cluster forms which encompasses all, bar

four, of the *R. meliloti* isolates and ten of the thirteen *Bradyrhizobium* isolates. This large group also contains a selection of isolates from the three biovars of the species *R. leguminosarum* but mainly those converged near to the origin. The remaining 31 strains, not incorporated into this large cluster, remain as individual, unlinked isolates at this level of similarity.

Three clusters form when the level of similarity examined is reduced to 75-80%. The largest of these groups links the ungrouped *Bradyrhizobium* and *R. meliloti* isolates together with nine isolates from the species *R. leguminosarum* (five biovar *trifolii*, two biovar *viciae* and two biovar *phaseoli*) to the cluster formed at the 80-100% level of similarity. The second largest group (at the 75-80% similarity level) contains seven *R. leguminosarum* biovar *viciae* isolates which lie in the lower right-hand segment of the plot. The third and smallest cluster groups three isolates of *R. leguminosarum* biovar *phaseoli* together. At this level of clustering six strains remain ungrouped. All 84 isolates are joined into one cluster at the 70-75% level of similarity.

As with the analysis of data from the results of amplification with SPH3 it is not possible to draw many definitive conclusions about the identification of particular species. However the spatial orientation of the two biovars of *R. leguminosarum* (biovars *trifolii* and *viciae*) suggest it may be possible to use this primed reaction to identify unknown isolates of these two biovars. This observation is very interesting as the two biovars are closely related and a method of discriminating between them would be useful. The results also indicate that with the use of the correct primer or primers for amplification it should be possible to differentiate between these two biovars. Welsh and McClelland (1991) used three primers in single primed amplification reactions and in a pairwise arrangement to achieve double primed reactions. They report that with these three primers it was possible to map polymorphisms in an inbred line of mice. By increasing the number of variables, i.e. primers, in this way one starts to increase the ability to discriminate between individuals. Carried to its extreme one would finish with a complement of primers covering the whole genome. It should therefore be possible to use various primers to

totally classify *Rhizobium* and *Bradyrhizobium* populations such as those used in this study.

In order to examine the relationship between the *R. leguminosarum* biovars *trifolii* and *viciae* the band data for these two biovars was examined separately from the other species. The results obtained following analysis of strains from *R. leguminosarum* biovars *trifolii* and *viciae* only are shown in Figure 4.6. A straight line is formed after this analysis. The *R. leguminosarum* biovars *trifolii* isolates, except two, can be seen to lie on the upper end of the line above the origin (0,0). Twelve of the *R. leguminosarum* biovar *viciae* strains are found in the lower end of the group with ordinate values of less than zero with respect to both axes. When Cluster Analysis is applied to these data (this is not shown on the plot for clarity) the apparent variation between the biovars *trifolii* and *viciae*, as seen by their spatial separation, no longer holds true. Small groups of up to six isolates, from one biovar only, are observed at a similarity level of 80% and over. However, at lower levels of similarity the groupings become mixed and all the strains eventually converge at a similarity level of 70%. This result does, however, tend to support the belief that these two biovars are highly related. When a comparison of all isolates from the species *R. leguminosarum* is made by adding the data for *R. leguminosarum* biovar *phaseoli* the ordinate plot shown in Figure 4.7 is produced. From this it can be seen that the *R. leguminosarum* biovars *trifolii* and *viciae* strains form a straight line similar to that in Figure 4.6. However the isolates from the two biovars are more intermingled in this case. The *R. leguminosarum* biovar *phaseoli* isolates, however, form a loose group off to the side of the line formed by the biovar *trifolii* and *viciae* isolates. This would indicate that this biovar (*phaseoli*) is not as related to the other two biovars as they are to each other. Overall these results reflect those found in Chapter 3 and reported by Segovia *et al.*, (1991) and Dooley *et al.*, (1993). It can be concluded, therefore, that the use of these two primers, in a double primed reaction, for the identification of individual biovars from a species is not as useful as it first appears. This may, however, be the

result of a small sample size (36 isolates) and as this is increased the level of detection may increase also.

Additional analysis:

In order to produce a single, large data set from two or more smaller sets the 1s and 0s were simply added together end to end. This, in effect, artificially increased the number of products being compared. Had the two original data sets been recalculated it may not have resulted in as large a number of products being examined. However it was felt that treating two RAPD products of the same molecular size, although produced in different primed reactions, as identical was not acceptable as they will have different termini sequences. This means they will have derived from different parts of the genome therefore cannot be regarded as identical DNA fragments.

The range of products observed when two sets of results are combined differs from those if a new RAPD reaction is carried out using the two primers in a double-primed reaction. This can be explained by using the first combination outlined below, i.e. that of SPH1 and SPH3. Following a RAPD reaction using the primer SPH1 alone all the bands produced will have SPH1 primer biding sites at their termini while those bands from a SPH3 primed reaction will have SPH3 primer sites at the end. Combining data from these two reactions results in an increase in the product number, but all products still have either a SPH1 or SPH3 primer sites at their termini. If, however, the two primers are combined in a double-primed RAPD reaction we would still expect an increase in the number of products as outlined above, but not all the products will have the same primer site at both ends. Some products will have an SPH1 binding site at one end and a SPH3 site at the other. Additionally the average band size decreases when two primers are used in a double primed RAPD reaction, however, when two data sets are combined this does not happen.

Combined data from SPH1 and SPH3 primed reactions:

When the data from amplification reactions primed with the primer SPH1 was combined with the results from SPH3 primed reactions the number of product under investigation totalled 105 (45 from SPH1 primed reactions and 60 from SPH3 primed reactions). The similarity matrix obtained following analysis of this data set is shown in Appendix 9.

Following analysis of this new data set the PCO plot shown in Figure 4.8 was produced. The results of analysis from a combination of SPH1 and SPH3 data shows an overall resemblance to that of the results from analysing SPH1 data alone. Spatial orientation within Figure 4.8 reveals that the *R. meliloti* isolates form two groups and that the *Bradyrhizobium* can also be distinguished, to a certain extent, from the remaining strains. The isolates of *R. leguminosarum* biovar *phaseoli* also form a distinct group but again the *R. leguminosarum* biovars *trifolii* and *viciae* cannot readily be distinguished from each other. When Cluster Analysis was applied to the data eighteen clusters were formed at the highest level (80-100%) of similarity. These clusters, except for two, are not shown on the plot for clarity. The two groups highlighted link four and two strains of *R. leguminosarum* biovar *phaseoli* respectively. These two groups are shown as they remain independent of the other groupings even at the lowest level of similarity examined in this study. The remaining clusters observed at the 80-100% level of similarity are mainly comprised of between two and nine isolates of the same species or biovar. When the 75-80% similarity level is examined ten clusters can be observed, the main one containing the isolates from *R. leguminosarum* biovars *trifolii* and *viciae* along with five *R. leguminosarum* biovar *phaseoli* isolates, three *Bradyrhizobium* and three *R. meliloti* isolates. There are also three groups containing *Bradyrhizobium* isolates only, two sets of *R. meliloti* only and three clusters formed by the linking of *R. leguminosarum* biovar *phaseoli* isolates only. The final cluster observed at this level of similarity contains two *R. leguminosarum* biovar *trifolii* isolates and one isolate each of *R. meliloti* and *R. leguminosarum* biovar *viciae*. This clustering pattern is similar to that obtained when SPH1 was analysed

alone (Chapter 3). From Figure 4.8 it can be seen that at the 70-75% level of similarity only four groups are in evidence; one contains *Bradyrhizobium* and *R. meliloti* isolates only, two contain *R. leguminosarum* biovar *phaseoli* isolates only and the fourth contains the *R. leguminosarum* biovars *trifolii* and *viciae* with a few isolates from the other species.

From comparisons of this plot to those produced by SPH1 and SPH3 alone it would appear that the effects caused by amplification by SPH1 have a greater influence on this combined data than those caused by SPH3. Data from SPH3 amplification alone shows no useful traits which could be used for classification studies and in this combined data set it appears to be only having a deleterious effect on the original data obtained from SPH1 amplification. This would indicate that in this case the combining of two data sets shows no marked increase in clarity of the results and so is of no practical use.

Combined data from SPH1 and SPH3+7 primed reactions:

The number of bands being considered when SPH1 and SPH3+7 data was combined totalled 130 and the similarity matrix produced following analysis of this data is shown in Appendix 10. The PCO plot obtained after analysis of combined data from SPH1 and SPH3+7 is shown in Figure 4.9. Like the data from the combination of SPH1 and SPH3 the strong influence of SPH1 data shows through in the general spatial layout of isolates. This would suggest that the combining of data in the way done here is not having a negative effect on the overall combined results, i.e. the larger number of products from SPH3+7 results (which is almost double that of SPH1) is not unduly skewing the analysis, but the quality of the data is having a greater effect. The species and biovar orientations on this plot are remarkably similar to those seen on the plot when SPH1 data only is analysed (Figure 3.2, Chapter 3). This suggests that the SPH1 data is having a strong influence on the combined data set being examined here. From observations of Figure 4.9 it can be seen that the *R. leguminosarum* biovar *phaseoli* strains form a loose group at the upper edge of the plot. The *R. meliloti*

strains form two clusters down the left margin of the plot with the *Bradyrhizobium* isolates laying along the bottom edge of the plot in a long drawn out shape. The *R. leguminosarum* biovars *viciae* and *trifolii* isolates are found mostly in the right hand region of the plot. However on closer examination of the spatial separation of the strains one major difference between the two PCO plots (Figures 3.2, Chapter 3 and Figure 4.9) can be seen. The strains of *R. leguminosarum* biovars *trifolii* and *viciae* appear to be more easily distinguishable, within their general cluster, with the combined data than they are with SPH1 data only. This suggests that the influence of the SPH3+7 data is having a greater effect on these biovars than the SPH1 data and thus combining the two data sets appears to be advantageous in this case. This observation holds true even when Cluster Analysis has been applied to the plot. This result would suggest that certain primers are better than others for the elucidation of variations between strains of a single species. By employing a larger range of primers it may therefore be possible to identify single individuals from a particular species. When two primers are used in double primed reactions there is usually an increase in the number of bands produced (Welsh and McClelland, 1991). This was also noted in the reactions when SPH3 and SPH7 were used together. Caetano-Anollés *et al.*, (1991) have reported that complex patterns (such as these double primed reactions) contain more information and so are of more use for genotyping. It may be that the increased band number and hence pattern complexity observed in double primed reactions allows for a more detailed examination of individuals. The use of a variety of primer combinations for *Rhizobium* studies should make it possible to identify individuals of the species *R. leguminosarum* biovars *trifolii* and *viciae*. The primers SPH1 and either SPH3 or SPH7, if used in a double primed reaction may be useful for individual strain identification as SPH1 appears to be very useful for classifying species whilst SPH3 and SPH7 are fairly useful for individual isolate identification within particular species.

At the 80-100% level of similarity fifteen clusters can be observed; two contain *Bradyrhizobium* isolates only, a third contains *R. meliloti* strains only and a fourth, larger group forms when four *Bradyrhizobium* strains cluster with ten *R. meliloti*

isolates. There are also five clusters with *R. leguminosarum* biovar *phaseoli* isolates only and one linking two *R. leguminosarum* biovar *phaseoli* isolates and two *R. meliloti* isolates. The *R. leguminosarum* biovar *viciae* isolates form four clusters which account for thirteen of the isolates, while a further four remain ungrouped. The final *R. leguminosarum* biovar *viciae* isolate links with a large group comprising seventeen of the eighteen *R. leguminosarum* biovar *trifolii* isolates along with two *Bradyrhizobium* and one *R. meliloti* isolates. The results of this Cluster Analysis seem to support the hypothesis that the individual influences of the SPH1 and SPH3+7 data produce a better overall separation of strains, especially with regard to the *R. leguminosarum* biovar *viciae* and *trifolii* isolates. The *R. leguminosarum* biovar *viciae* and *trifolii* isolates were analysed separately as described above. However results similar to those found earlier were observed and so are not shown here. When the 75-80% level of similarity is examined five clusters are present. One collection contains primarily *R. leguminosarum* biovar *phaseoli* isolates along with two *R. meliloti* isolates. Two groups contain three isolates of *R. leguminosarum* biovar *viciae* each, one is formed by the clustering of six *Bradyrhizobium* strains from three previously formed groups. The remaining isolates form the final cluster at this level of analysis (not shown for clarity). A single group of all 84 isolates is present at the 70-75% level of similarity. This early clustering of all isolates makes it impossible to draw any meaningful conclusions from these results.

Combined data from SPH3 and SPH3+7 primed reactions:

Figure 4.10 shows the PCO plot produced following analysis of combined data from SPH3 primed and SPH3+7 primed amplification reactions. Combining data from these two amplification reactions resulted in 145 products being compared. This resulted in the production of the similarity matrix shown in Appendix 11. From observations of this plot it can be seen that certain areas appear to contain strains from specific species or biovars only. All the *R. leguminosarum* biovar *viciae*, except four, fall on the left of the graph within the region of less than -0.1 with respect to ordinate

1. These strains are bounded within the area -0.2 to 0.2 with respect to ordinate 2. Within this region there are only three strains which are not *R. leguminosarum* biovar *viciae* strains; two *R. leguminosarum* biovar *trifolii* strains and one *Bradyrhizobium* strain. The *R. leguminosarum* biovar *phaseoli* strains, with the exception of three isolates, fall within the region of the plot bounded by greater than -0.05 and less than 0.21 with respect to ordinate 1 and greater than 0.1 with respect to ordinate 2. The *R. meliloti* strains appear to be mainly located in the lower right-hand quadrant of the plot along with the *Bradyrhizobium* isolates. The *R. leguminosarum* biovar *trifolii* strains occupy a central region of the graph between -0.1 and 0.1 with respect to ordinate 1 and (except for two strains) greater than 0.0 with respect to ordinate 2. From simple spatial orientation it would, therefore, appear that the combination of these two data sets is advantageous for strain identification. However, when Cluster Analysis is applied this apparent spatial separation dissolves. Cluster Analysis has not been shown on the plot for clarity, however, at the 80-100% level of similarity 21 clusters form within the plot. These are composed of between two and eight strains in size. The number of clusters is reduced to ten at the 75-80% level of similarity and finally six at the lowest level of similarity considered (70-75%). The groups formed, especially at the 75-80% and lower levels of similarity, pay little consideration to species or biovars and are mostly formed from several strains types.

It was concluded that although the spatial separation seen with this combined data set appeared useful for strain identification it is, in fact, of limited use. However, the marked separation of the *R. leguminosarum* biovars *trifolii* and *viciae* suggests that the use of a triple combination data set may result in a better final plot where it may be possible to classify all the *Rhizobium* and *Bradyrhizobium* strains used in this study.

Combined data from SPH1 and SPH3 and SPH3+7 primed reactions:

The combining of three data sets results in the largest number of products being examined. This amounts to 190 amplified products, which gave rise to the similarity matrix shown in Appendix 12. Following analysis of this data the PCO plot shown in Figure 4.11 was produced. From simple observations of the spatial distribution of the strains the plot appears quite similar to that when SPH1 data was analysed alone. Within the group containing strains of *R. leguminosarum* biovar *viciae* and biovar *trifolii*, small clusters of isolates, from the two biovars, appear to be emerging. When Cluster Analysis is applied to the plot the strains remain in fairly consistent groupings especially at the higher, 80-100%, level of similarity used. At this level of analysis nineteen clusters are in evidence. There are four groups of *Bradyrhizobium* strains only, one consisting of three strains and three of two isolates each. A single *Bradyrhizobium* isolate links with two *R. meliloti* strains to form a fifth cluster, while there are also two groups containing *R. meliloti* isolates only. One of these comprises nine strains and the other three. Three sets of *R. leguminosarum* biovar *phaseoli* can be seen, one of five isolates and the other two composed of two strains each. The eleventh cluster comprises two *R. leguminosarum* biovar *phaseoli* isolates and a *R. leguminosarum* biovar *trifolii* strain. The last *R. leguminosarum* biovar *phaseoli* strain to cluster at this level of similarity links with a *R. leguminosarum* biovar *trifolii* strain to form a two strain group. The remaining *R. leguminosarum* biovar *phaseoli* strains do not group at this level of similarity. The final seven clusters formed at the 80-100% level of similarity are comprised of *R. leguminosarum* biovar *trifolii* and *R. leguminosarum* biovar *viciae* strains. There are three groups of *R. leguminosarum* biovar *viciae* strains only, one each of two, three and four isolates. There is one group of two *R. leguminosarum* biovar *trifolii* strains only. The final, and largest, cluster at this level of similarity comprises eleven *R. leguminosarum* biovar *trifolii* and three *R. leguminosarum* biovar *viciae* isolates.

When the 75-80% level of similarity is considered only eight clusters remain. Two of these are unlinked groups (one *R. leguminosarum* biovar *phaseoli* and one

Bradyrhizobium) formed at the 80-100% similarity level. The third cluster is formed when the two groups of *R. meliloti* strains join with the previously formed cluster comprising two *R. meliloti* and one *Bradyrhizobium* strains. A large cluster is formed by the clustering of all (except three) of the *R. leguminosarum* biovar *phaseoli* strains with two *R. leguminosarum* biovar *trifolii* isolates. Two groups, one of seven and one of four, strains of *R. leguminosarum* biovar *viciae* are formed at this level of similarity. A smaller group forms when a single *R. meliloti* isolate links with two *R. leguminosarum* biovar *trifolii* strains. The final group seen comprises a mixture of seven *Bradyrhizobium*, fifteen *R. leguminosarum* biovar *trifolii* and seven *R. leguminosarum* biovar *viciae* strains.

Two large clusters remain at the 70-75% level of similarity. The smaller group comprises the large *R. meliloti* cluster formed at the 75-80% level and the three *Bradyrhizobium* strains which grouped at the 80-100% level of similarity. The remaining isolates form the second cluster.

The results obtained from combining all three data sets were not as clear as was hoped although at the 75-80% level of similarity it was possible to distinguish a proportion (ten) of the *R. leguminosarum* biovar *viciae* strains from the *R. leguminosarum* biovar *trifolii* strains. However the clustering of the *R. leguminosarum* biovar *trifolii* isolates with the *Bradyrhizobium* isolates at this level somewhat nullifies the result.

Conclusions:

From an overall examination of all the results outlined here it can be concluded that (i) the results from analysis of SPH1 primed amplification reactions provides the best means of classifying the *Rhizobium* and *Bradyrhizobium* strains used in this study; (ii) the use of primers SPH3 and 7 in a double primed reaction appears to confirm findings made with primer SPH1 alone and reported in Chapter 1. (iii) Combined data from SPH1 and SPH3 and 7 reactions appears to provide a fairly accurate method for distinguishing between *R. leguminosarum* biovar *viciae* and *R. leguminosarum* biovar

trifolii isolates. However, when these two biovars are analysed alone the accuracy is not so well maintained, with results being similar to those when SPH3 and 7 data only was analysed (Figure 4.6). (iv) Results from SPH3+7 double primed amplification reactions shows that although *R. leguminosarum* biovars *trifolii* and *viciae* are highly related there are discernible differences between them which could be exploited to allow total classification of this species. (v) SPH3 data is of little use for classifying the *Rhizobium* and *Bradyrhizobium* strains used in this study. Its use in combined data analysis even has a greater detrimental than beneficial effect on the overall analysis; (vi) Combining data sets by simply adding them together, which in effect increases the number of amplified products, does not appear (except in the case of SPH3 data) to unduly effect the analysis. (vii) The SPH1 data appears to have the greatest influence on the overall outcome when combined data is analysed. This occurs even though the number of products observed following amplification with the primer SPH1 alone is less than the number of products seen following amplification with the other primers (except SPH7). (viii) The effect of combining data with that from SPH1 primed amplifications is to increase the resolving power, especially between isolates of *R. leguminosarum* biovar *trifolii* and biovar *viciae*. This would indicate that although primer SPH1 is useful for species differentiation it may not be as useful as some other primers for classifying closely related biovars of the same species. (ix) Finally it can be concluded that it would appear that with the use of the correct primers in the correct combinations it should be possible to classify any *Rhizobium* or *Bradyrhizobium* isolate into its specific species or biovar.

Tables and Figures

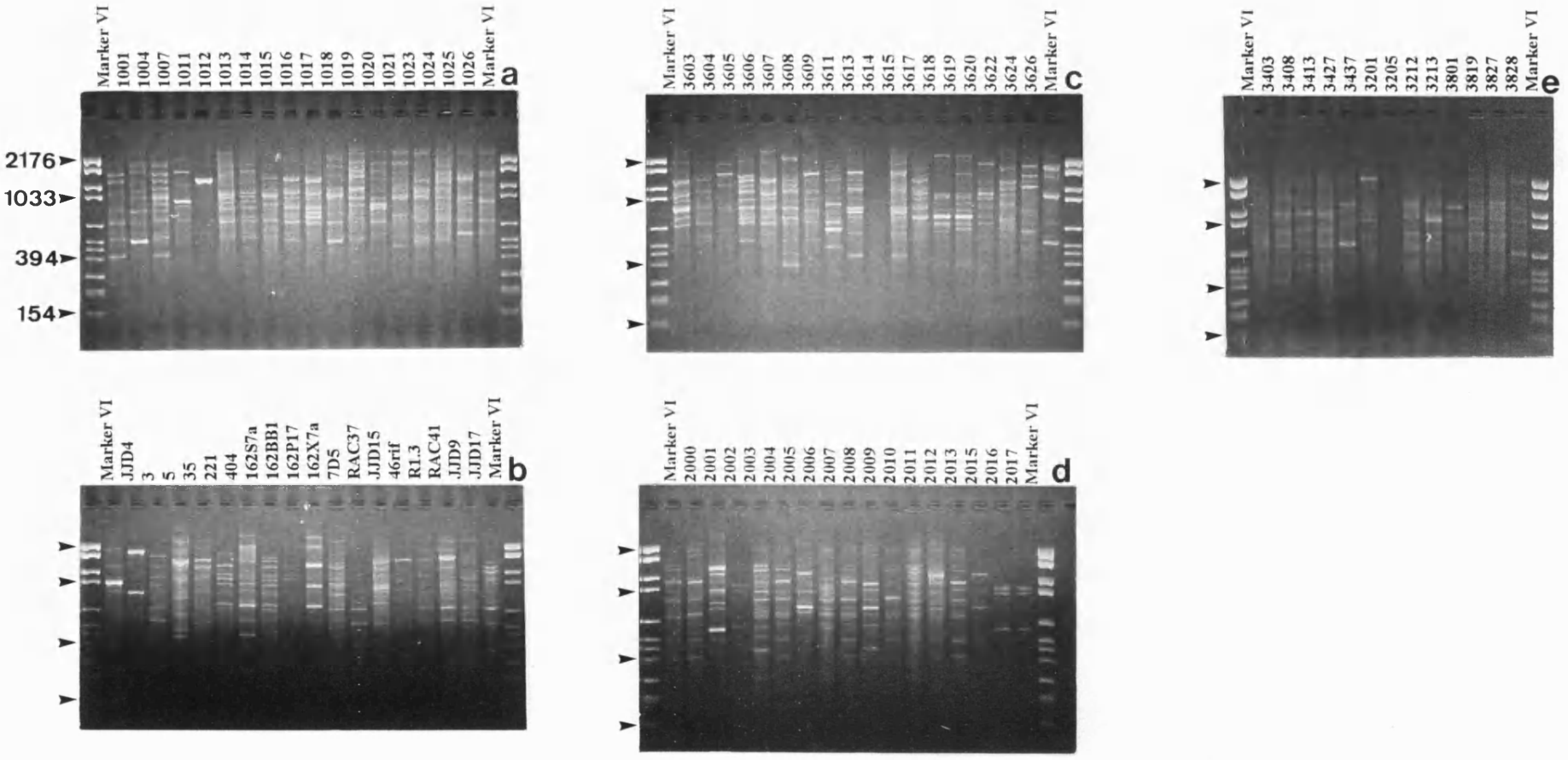


Figure 4.1: RAPD fingerprints produced using the primer SPH3. Patterns were obtained from strains of *R. leguminosarum* biovar *viciae* (a), biovar *trifolii* (b), biovar *phaseoli* (c), *R. meliloti* (d) and *Bradyrhizobium* (e). Strain names are marked above each lane. Marker VI (B. Mannheim) was added to the outer lanes. Marker sizes (arrowed) are in bp.

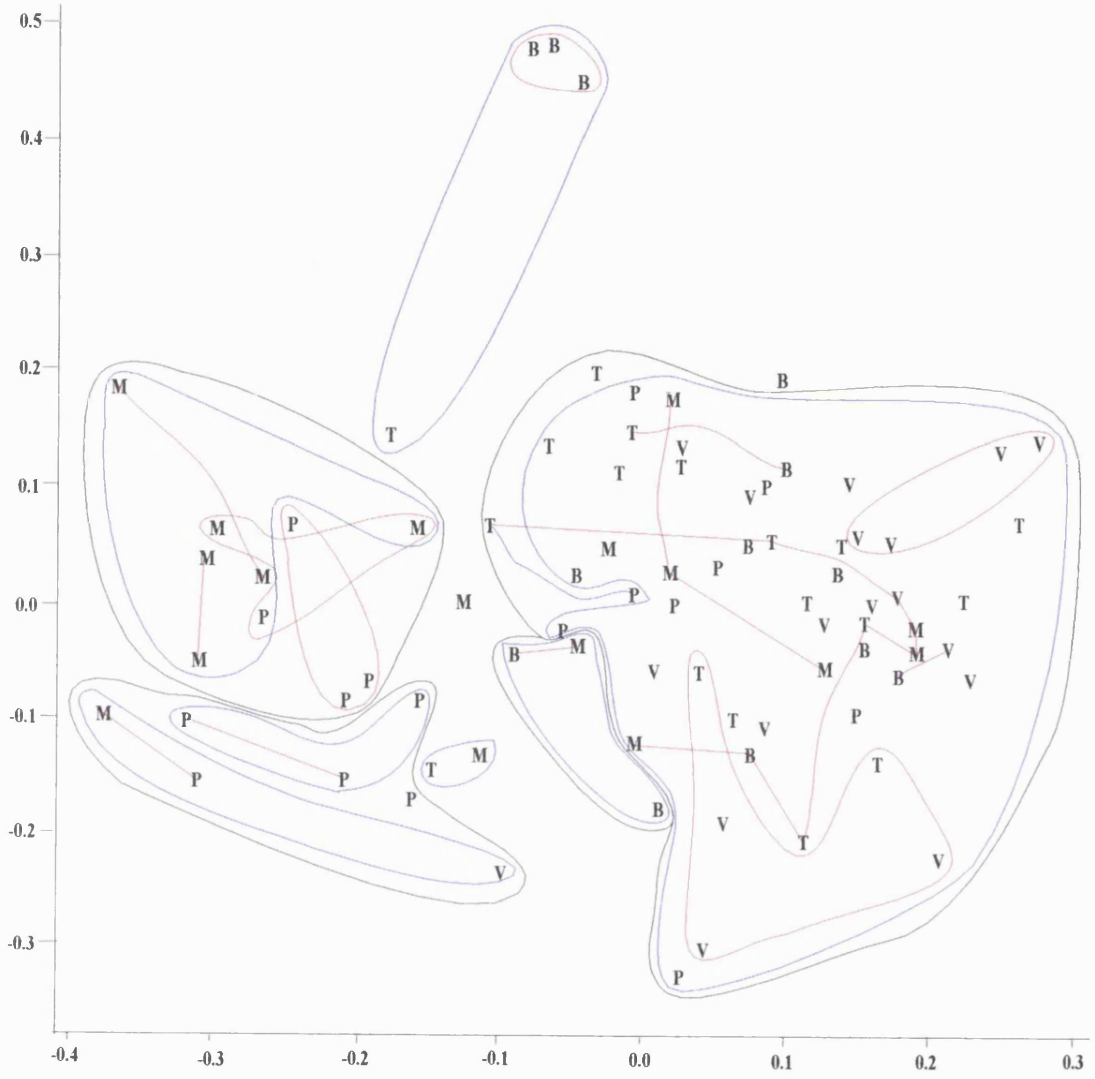


Figure 4.2: Principal Coordinate plot of ordinates 1 and 2 produced following analysis of data from DNA amplification with primer SPH3. Data was matched using Simple Matching. Isolates can be identified as *R. leguminosarum* bv *trifolii* (T), bv *viciae* (V), bv *phaseoli* (P), *R. meliloti* (M) and *Bradyrhizobium* (B). Isolates were grouped according to the results of Cluster Analysis which has been applied to the PCO plot using three levels of clustering; 70-75 % (——), 75-80 % (——) and 80-100 % (——).

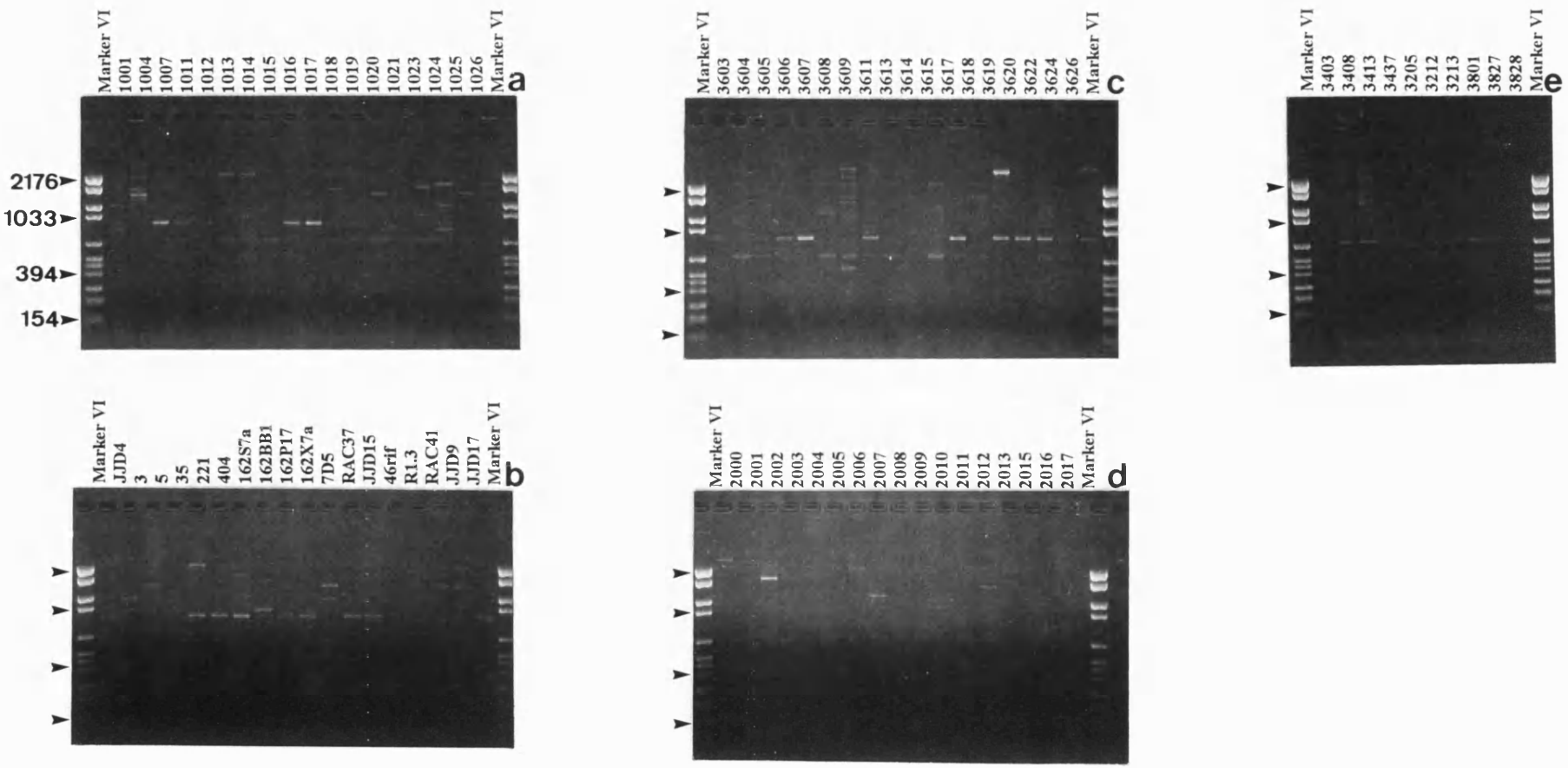


Figure 4.3: RAPD fingerprints produced using the primer SPH7 alone. Patterns were obtained from strains of *R. leguminosarum* biovar *viciae* (a), biovar *trifolii* (b), biovar *phaseoli* (c), *R. meliloti* (d) and *Bradyrhizobium* (e). Strain names are marked above each lane. Marker VI (B. Mannheim) was added to the outer lanes. Marker sizes (arrowed) are in bp.

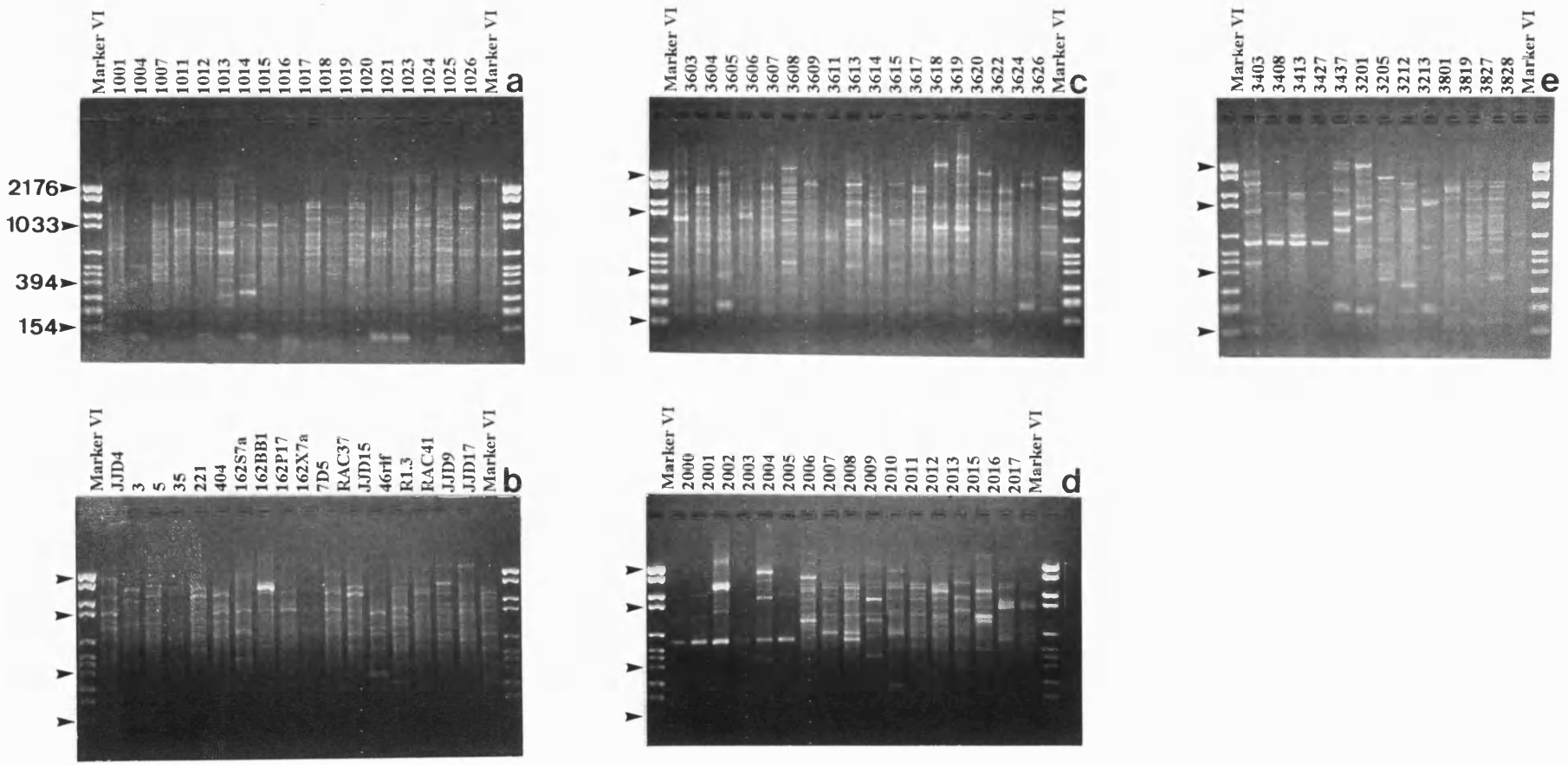


Figure 4.4: RAPD fingerprints produced using the primers SPH3 and SPH7 in a double primer reaction. Patterns were obtained from strains of *R. leguminosarum* biovar *viciae* (a), biovar *trifolii* (b), biovar *phaseoli* (c), *R. meliloti* (d) and *Bradyrhizobium* (e). Strain names are marked above each lane. Marker VI (B. Mannheim) was added to the outer lanes. Marker sizes (arrowed) are in bp.

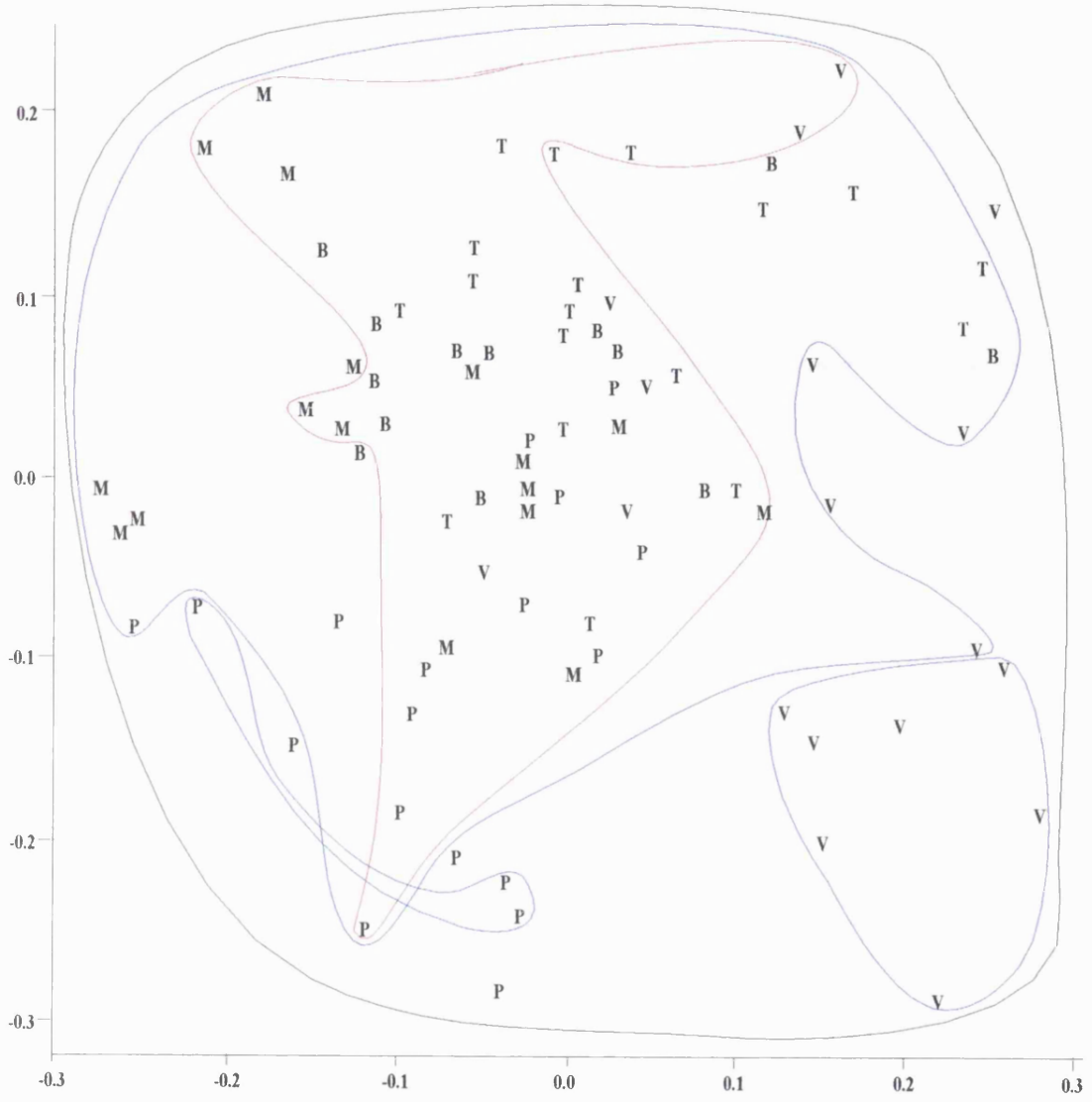


Figure 4.5: Principal Coordinate plot of ordinates 1 and 2 produced following analysis of data from DNA amplification with primers SPH3 and SPH7. Data was matched using Simple Matching. Isolates can be identified as *R. leguminosarum* bv *trifolii* (T), bv *viciae* (V), bv *phaseoli* (P), *R. meliloti* (M) and *Bradyrhizobium* (B). Isolates were grouped according to the results of Cluster Analysis which has been applied to the PCO plot using three levels of clustering; 70-75 % (——), 75-80 % (——) and 80-100 % (——).

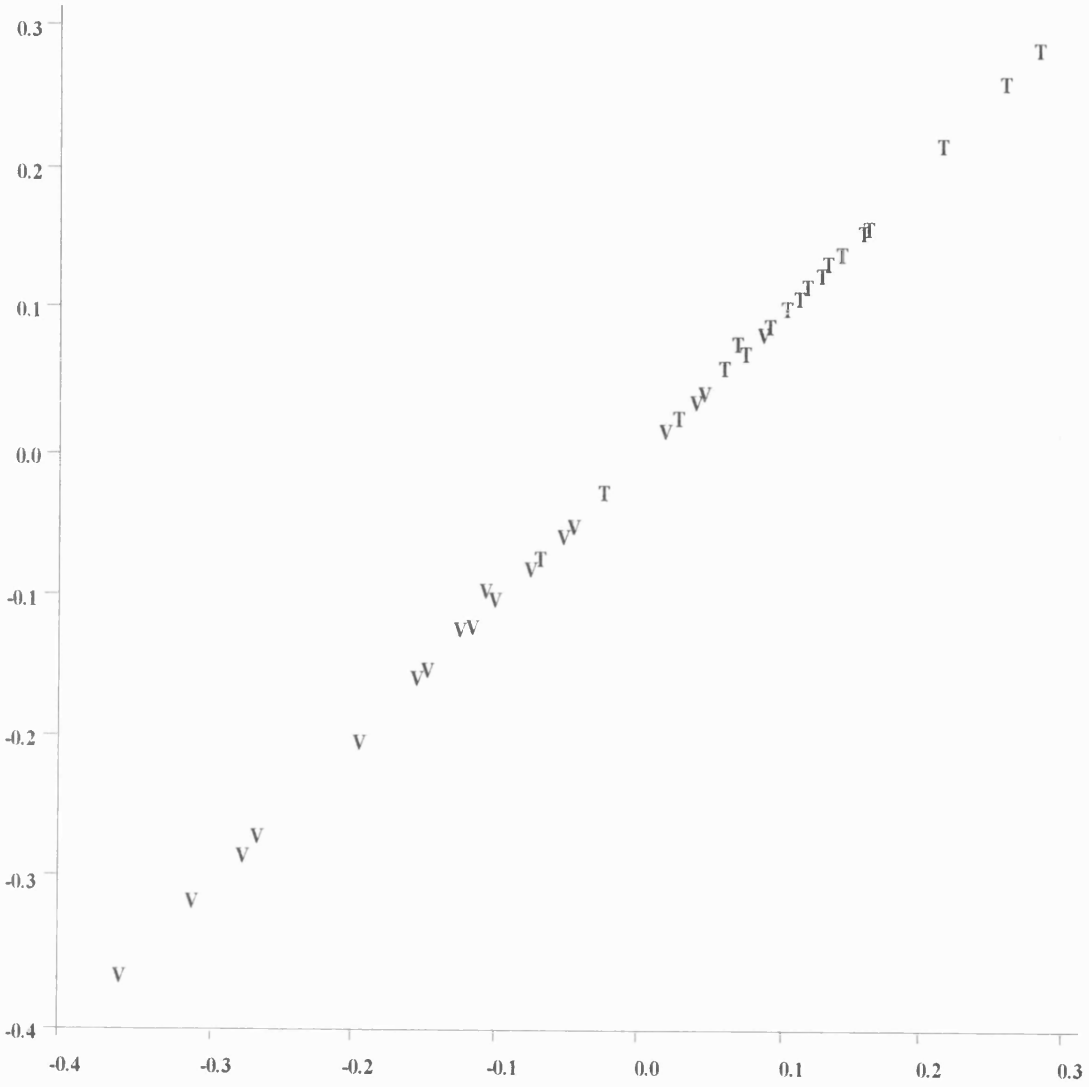


Figure 4.6: Principal Coordinate plot of ordinates 1 and 2 produced following analysis of data from DNA amplification with primers SPH3 and SPH7. Data was matched using Simple Matching. Isolates can be identified as either *R. leguminosarum* bv *trifolii* (T) or bv *viciae* (V).

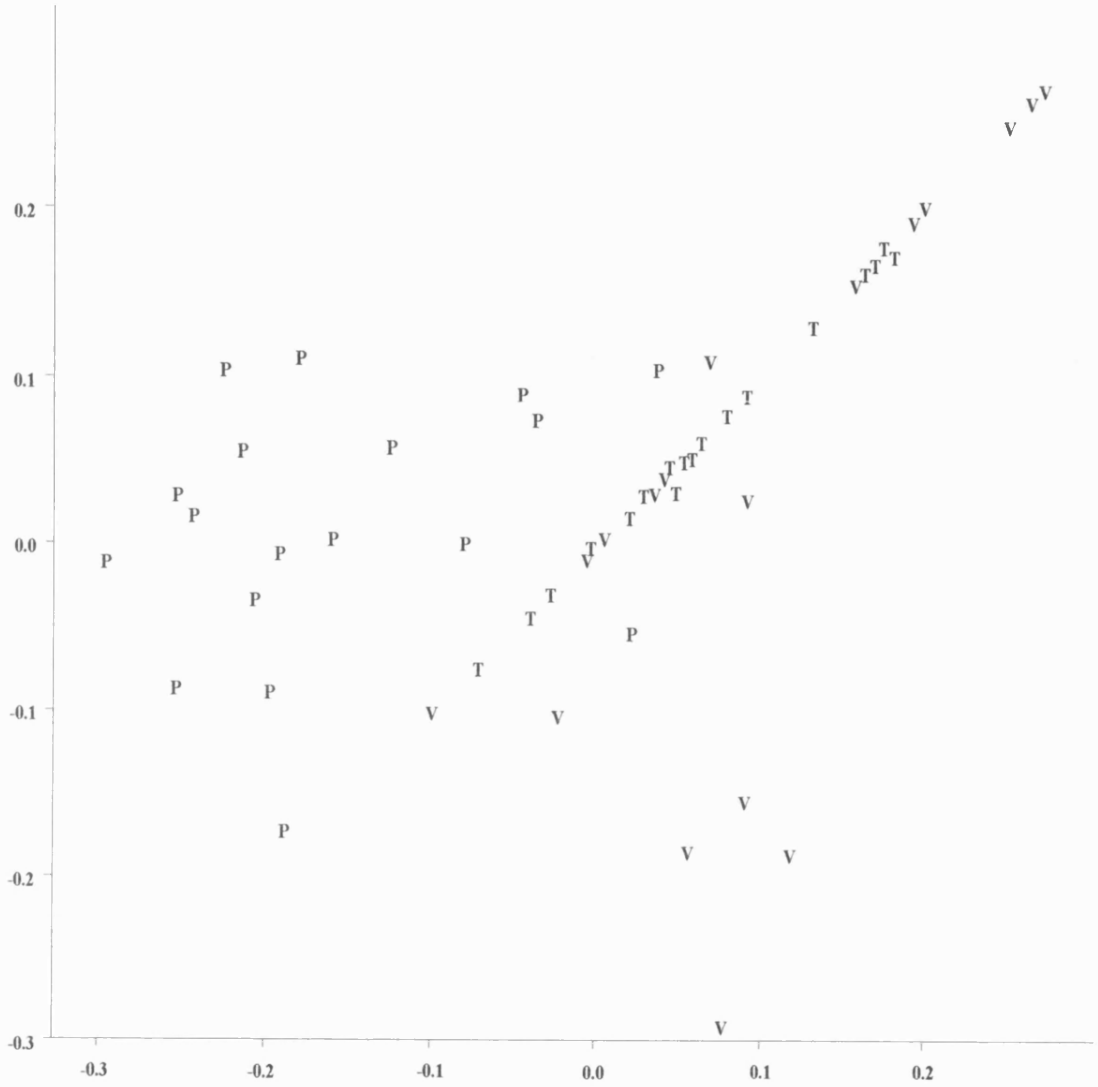


Figure 4.7: Principal Coordinate plot of ordinates 1 and 2 produced following analysis of data from DNA amplification with primers SPH3 and SPH7. Data was matched using Simple Matching. Isolates can be identified as *R. leguminosarum* bv *trifolii* (T), bv *viciae* (V) or bv *phaseoli* (P).

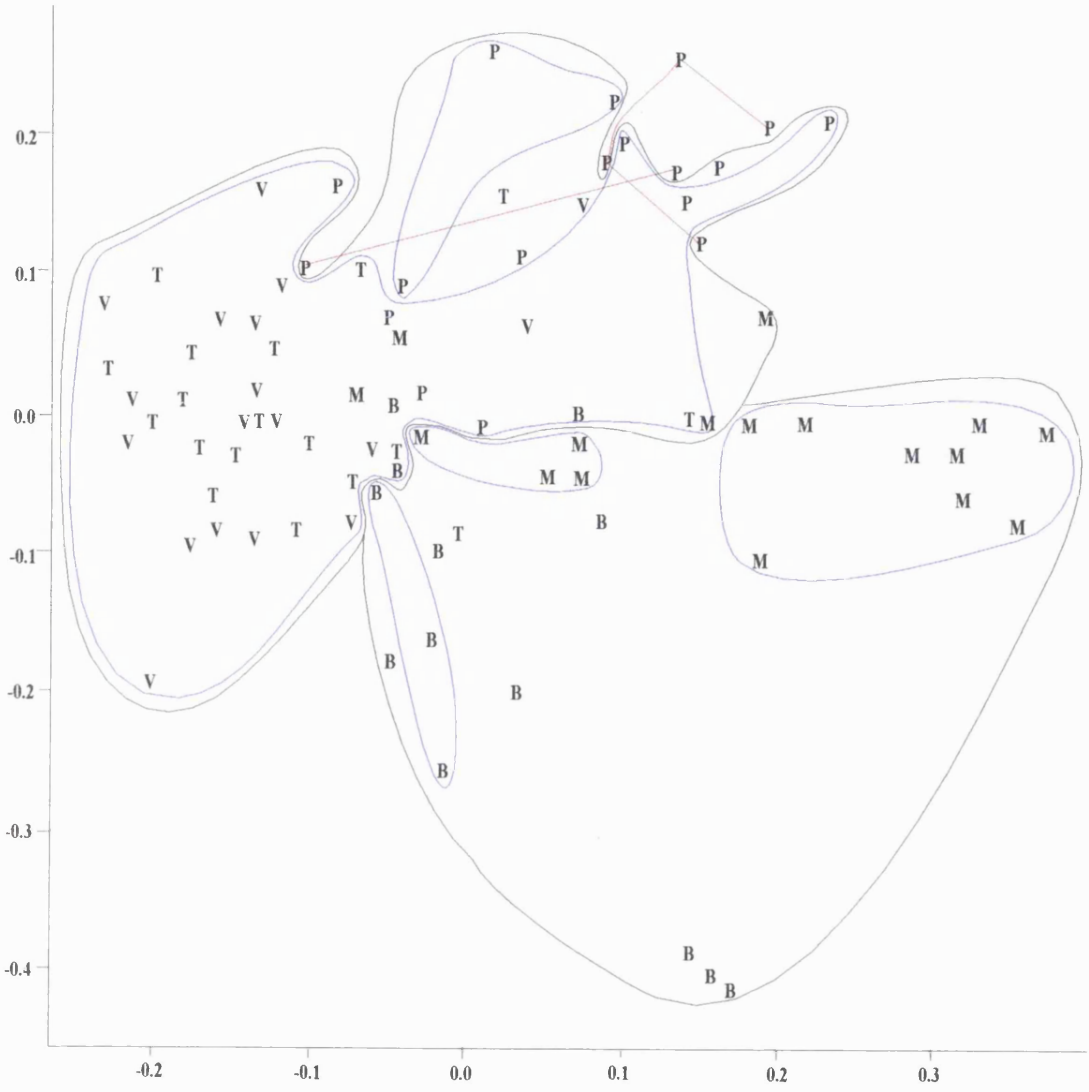


Figure 4.8: Principal Coordinate plot of ordinates 1 and 2 produced following analysis of combined data. The data sets were obtained following DNA amplification with primer SPH1 and amplification with primer SPH3. Data was matched using Simple Matching. Isolates can be identified as *R. leguminosarum* bv *trifolii* (T), bv *viciae* (V), bv *phaseoli* (P), *R. meliloti* (M) and *Bradyrhizobium* (B). Isolates were grouped according to the results of Cluster Analysis which has been applied to the PCO plot using three levels of clustering; 70-75 % (——), 75-80 % (——) and 80-100 % (——).

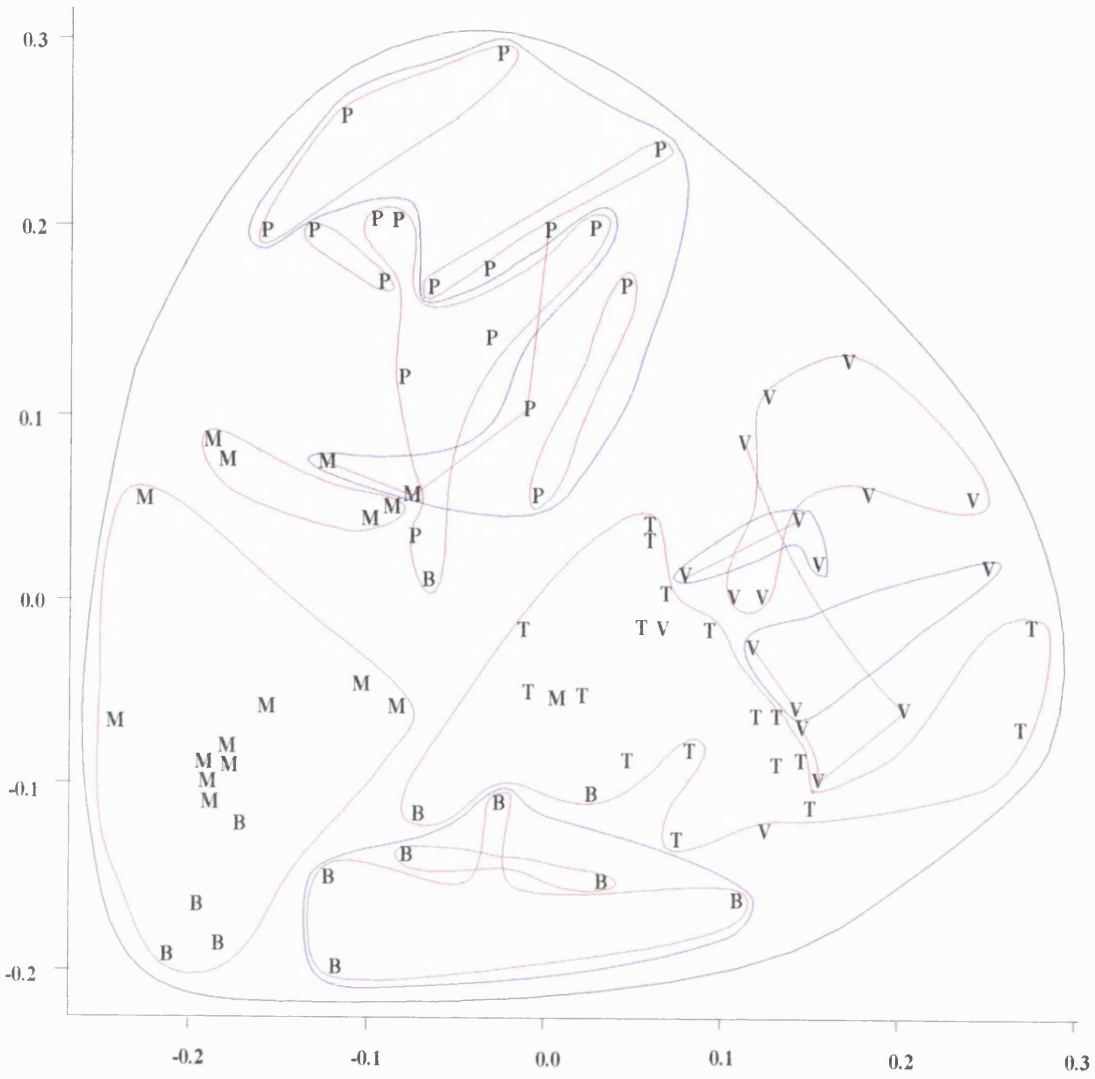


Figure 4.9: Principal Coordinate plot of ordinates 1 and 2 produced following analysis of combined data. The data sets were obtained following DNA amplification with primer SPH1 and amplification with primers SPH3 and SPH7. Data was matched using Simple Matching. Isolates can be identified as *R. leguminosarum* bv *trifolii* (T), bv *viciae* (V), bv *phaseoli* (P), *R. meliloti* (M) and *Bradyrhizobium* (B). Isolates were grouped according to the results of Cluster Analysis which has been applied to the PCO plot using three levels of clustering; 70-75 % (——), 75-80 % (——) and 80-100 % (——).

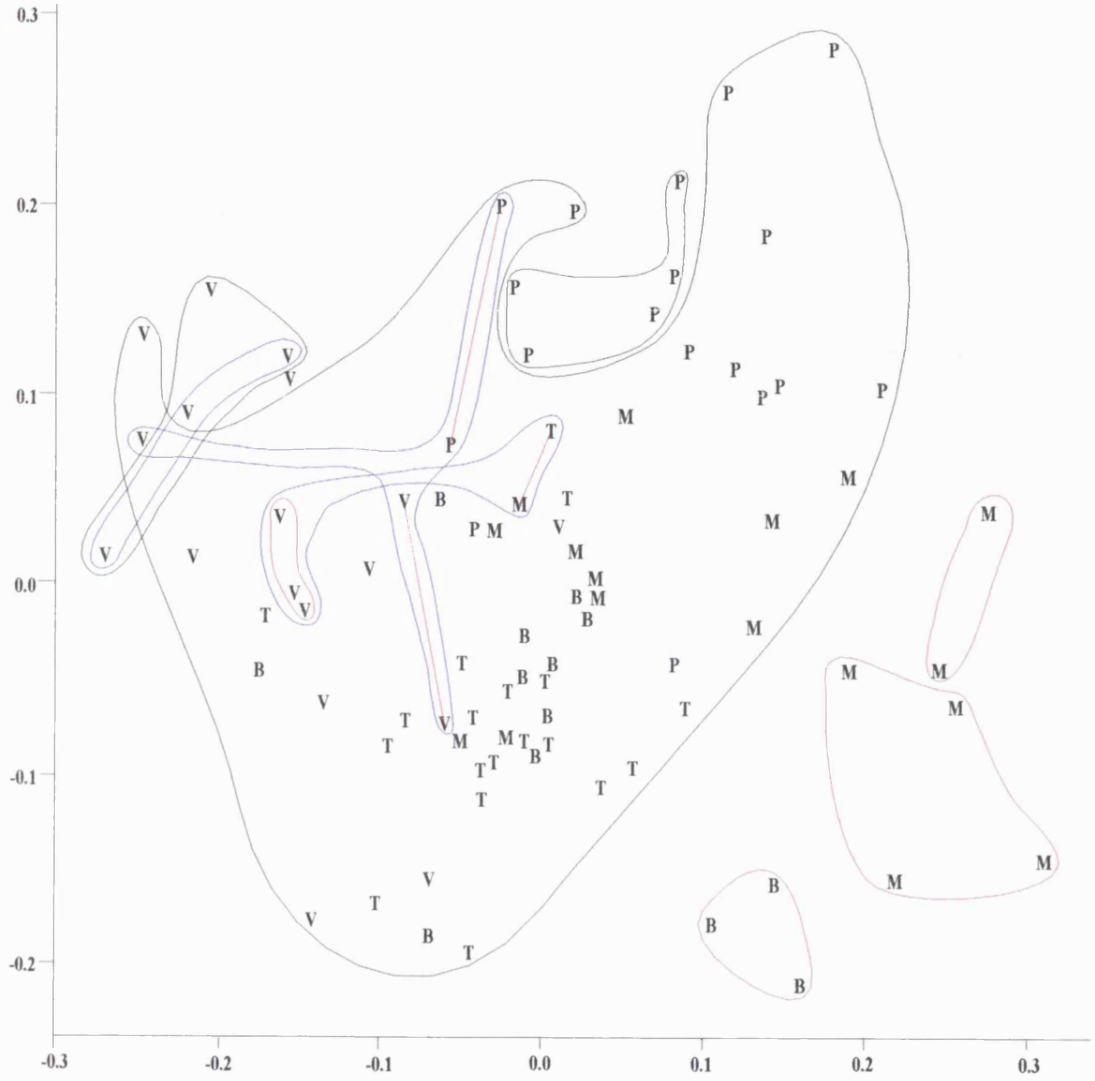


Figure 4.10: Principal Coordinate plot of ordinates 1 and 2 produced following analysis of combined data. The data sets were obtained following DNA amplification with primer SPH3 and amplification with primers SPH3 and SPH7. Data was matched using Simple Matching. Isolates can be identified as *R. leguminosarum* bv *trifolii* (T), bv *viciae* (V), bv *phaseoli* (P), *R. meliloti* (M) and *Bradyrhizobium* (B). Isolates were grouped according to the results of Cluster Analysis which has been applied to the PCO plot using three levels of clustering; 70-75 % (——), 75-80 % (——) and 80-100 % (——).

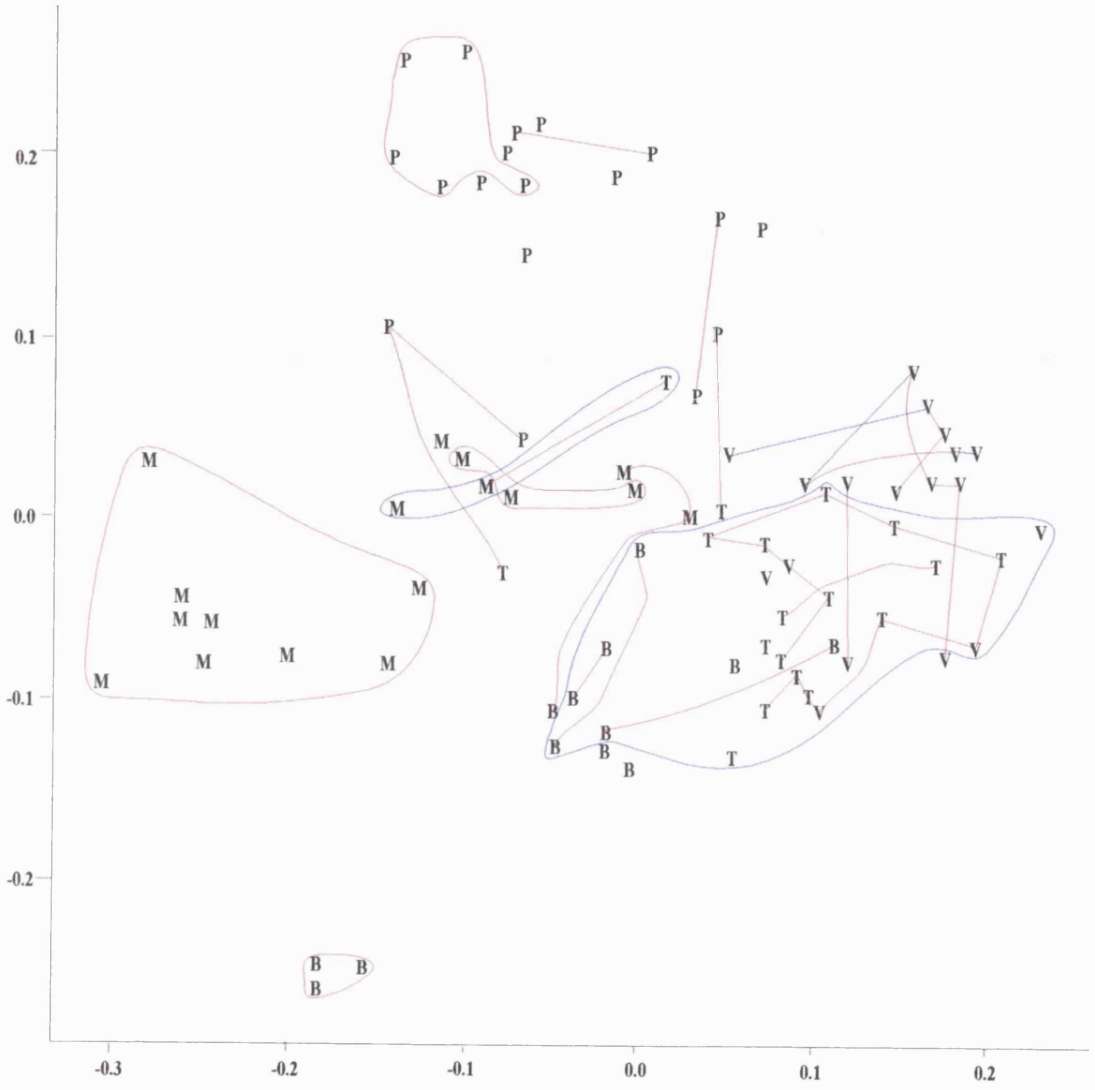


Figure 4.11: Principal Coordinate plot of ordinates 1 and 2 produced following analysis of combined data. The data sets were obtained following DNA amplification with primer SPH1, primer SPH3 and primers SPH3 and SPH7. Data was matched using Simple Matching. Isolates can be identified as *R. leguminosarum* bv *trifolii* (T), bv *viciae* (V), bv *phaseoli* (P), *R. meliloti* (M) and *Bradyrhizobium* (B). Isolates were grouped according to the results of Cluster Analysis which has been applied to the PCO plot using three levels of clustering; 70-75 % (——), 75-80 % (——) and 80-100 % (——).

An Assessment of the Specificity of Potential Species-Specific, RAPD-Derived Probes.

Material from this chapter has been included in the paper:

DOOLEY J.J. & HARRISON S.P. A Direct Method to Construct DNA Probes From RAPD Profiles. Submitted to FEMS Letters.

Abstract

Results of hybridising RAPD-derived probes back to RAPD profiles, fixed to nylon membranes, revealed that probes V1D and T1E, which were derived from bands found to be common to isolates of *R. leguminosarum* biovars *viciae* and *trifolii* respectively, are highly related. This observation suggests that the bands may be from identical parts of their respective genomes. However, a size difference of 40 bp does exist between the probes. This would indicate that there is some detectable variation between the two biovars. These probes do, however, display species-specificity in that they fail to hybridise with isolates not of the species *R. leguminosarum*. However, they also fail to hybridise with isolates of *R. leguminosarum* bv *phaseoli* which suggests that *R. leguminosarum* bv *phaseoli* may not be as related to the other biovars (*trifolii* and *viciae*), of the species *R. leguminosarum*, as these biovars are to each other. The three *R. meliloti*-derived probes examined show varying degrees of cross-reactivity with strains from *R. leguminosarum* bv *phaseoli* but reveal no cross-hybridisation to the remaining *R. leguminosarum* isolates. One of these probes, M1B, reveals quite high levels of cross-hybridisation, which indicates a high degree of similarity between isolates from *R. meliloti* and *R. leguminosarum* bv *phaseoli*. This again indicates a weaker relationship between *R. leguminosarum* bv *phaseoli* and the other biovars of the species *R. leguminosarum*.

Introduction

DNA hybridisation has been used for the identification of bacterial species, including *Rhizobium* and *Bradyrhizobium*, and is described more fully in Chapter 1. Total genomic DNA probes have been used to identify *R. loti* and *Bradyrhizobium* (Cooper *et al.*, 1987) and *R. leguminosarum* bv *trifolii* (Hodgson and Roberts, 1983). Reports have been made of the use of specific DNA probes for studies on the three biovars of *R. leguminosarum*. Demezas *et al.*, (1991) found that RFLP analysis, obtained using chromosomal probes, grouped strains into clusters which reflected results obtained from isoenzyme analysis. They also found that Sym plasmid probes where not as useful for classification. Schofield *et al.*, (1987) employed both Sym and chromosomal probes to study a population of *R. leguminosarum* bv *trifolii* isolates. Their results indicated that there was evidence for genetic exchange of Sym plasmids between strains. This was in contrast to observations made by Engvild *et al.*, (1990) who found limited genetic exchange, under natural conditions, among isolates of *R. leguminosarum* bv *viciae*. Laguerre *et al.*, (1992b) used two chromosomal probes and a Sym plasmid (*nod* gene) probe to study *R. leguminosarum* bv *viciae* isolates. They also found predominant plasmid hybridisation patterns within certain groups and a linkage between plasmid and chromosomal DNA hybridisation patterns, which suggested limited genetic exchange between strains. However, variations to these groups does suggest genetic exchange occurs. Specific probes have also been employed to study other species. Bjourson and Cooper, (1988) used subtraction-hybridisation to develop *R. loti* strain-specific probes, while Wheatcroft and Watson, (1987, 1988b) and Minamisawa *et al.*, (1992) used IS and RS as probes to study *R. meliloti* and *B. japonicum* respectively. Specific DNA probes were also employed by Streit *et al.*, (1993) who used subtraction-hybridisation and DNA amplification techniques to develop a probe capable of detecting *R. leguminosarum* bv *phaseoli* and *R. tropici* strains only. By using total genomic DNA from eight subtracter species (all

of which were of the family Rhizobiaceae) they identified sequences which were specific to isolates of *R. leguminosarum* bv *phaseoli* and *R. tropici* only.

The use of RAPDs for the generation of species-specific probes has been described by Dobrowolski & O'Brien (1993) who employed RAPDs to obtain probes for the pathogenic fungi *Phytophthora cinnamomi*. These researchers cloned potential probes into pUC18 for screening purposes. A report by Bjourson and Cooper (1992) reveals that by using the technique of band stabbing it is possible to produce probes without the need to clone the RAPD products into a plasmid. This technique overcomes the problems of cloning RAPD products and greatly reduces the time needed to screen such probes.

By using the technique of band stabbing (Bjourson and Cooper, 1992) it was hoped to ascertain if it was possible to generate species-specific probes from a series of bands common to the majority of isolates of *R. leguminosarum* biovars *trifolii* and *viciae* and *R. meliloti* which have been discussed earlier (Chapter 3). It was also hoped to determine if band stabbing could be used to generate non-isotopic DNA probes by labelling the DNA with digoxigenin.

Materials and Methods

Potential species-specific probes employed in this study were isolated from RAPD profiles produced using primer SPH1. These profiles have been described previously (Chapter 3; Dooley *et al.*, 1993). The bands of interest were isolated from their respective strains using the band-stab technique, as described in Chapter 2. The RAPD fingerprints and selected bands are shown Figure 5.1 and the probes are detailed in Table 5.1.

Results and Discussion

By combining the technique of band stabbing with DNA amplification we have found it is possible to produce digoxigenin labelled probes. This method eliminated the need to clone the RAPD product prior to labelling with Dig-dUTP. The incorporation of Dig-dUTP into the probe was confirmed by comparing the migration of probe DNA against its analogous parental DNA band through an agarose gel. The probe DNA, having the digoxigenin group on the side, runs at a slower rate than the equally sized non-Dig-labelled DNA band therefore a slower migration through the gel, by the probe, is indicative of successful labelling. The probes employed in this study are detailed in Table 5.1, which shows their species-specificity, derivative strain, molecular size and the frequency in the population, which has been calculated on presence of the derivative band within each species.

The results obtained when RAPD profiles were probed with the RAPD derived probes are shown in Figures 5.2 to 5.5. From a general observation of the results it was noted that all the probes appear to be fairly species-specific with low, if any, cross-reactivity to the other species being observed. It should be noted that the outer lanes from the 1.5 % RAPD gels failed to blot to the nylon filters fully so that the presence of some bands (particularly *R. meliloti* isolate *Rm*2000) are difficult to see. None of the probes used in this study hybridised to the isolates of the genus *Bradyrhizobium*. Cooper *et al.*, (1987) also report finding no intergeneric hybridisation between *R. loti* and *Bradyrhizobium* species when total DNA probes were used.

Probe M1A:

The probe M1A shows no hybridisation with RAPD profiles of strains of *R. leguminosarum* bv *trifolii* or *R. leguminosarum* bv *viciae* and so would appear species-specific. However when probed onto the *R. leguminosarum* bv *phaseoli* and *R. meliloti* the results in Figure 5.2 were produced. From this figure three faint bands of sizes 820, 630 and 540 bp can be seen in *R. leguminosarum* bv *phaseoli* isolate

*Rp*3604. The 540 bp band is also present in strains *Rp*3607 and *Rp*3622 and the 630 bp band can be observed in isolates *Rp*3608, *Rp*3618 and *Rp*3619. This 630 bp band displays a relatively strong signal in isolate *Rp*3618. The occurrence of cross-reactivity between these two species (*R. leguminosarum* bv *phaseoli* and *R. meliloti*) indicates a degree of common DNA is shared by isolates of both species. This suggests that the *R. leguminosarum* bv *phaseoli* isolates may have a common ancestry with strains from the species *R. meliloti* and that although they now nodulate beans they still maintain a percentage of the original genome DNA. This may be the case as previous reports have suggested that *R. leguminosarum* bv *phaseoli* may not be a homogeneous group, but may be comprised of a varied selection of phenotypically similar isolates (Chapter 3; Pinero *et al.*, 1988; Dooley *et al.*, 1993). Martinez-Romero *et al.*, (1991) have also proposed a new species, *R. tropici*, be formed from those strains formerly known as *R. leguminosarum* bv *phaseoli* Type II. There is however, insufficient data on the strains used here to permit them to be assigned to either the species *R. tropici* or *R. leguminosarum* bv *phaseoli*. It is, therefore, not possible to determine which of these species may have closer links with *R. meliloti*.

When M1A is hybridised back to the RAPD profiles obtained from *R. meliloti* isolates two, possibly polymorphic bands of size 720 bp and 650 bp (the size of the M1A) are seen. These bands appear as doublets in isolates *Rm*2000, *Rm*2003, *Rm*2005, and *Rm*2008. The larger band appears alone in the isolates *Rm*2011 and *Rm*2013 and the smaller band in *Rm*2001, *Rm*2002, *Rm*2006, *Rm*2007, *Rm*2010 and *Rm*2015. Isolate *Rm*2003 appears to contain both alleles at equal concentration whereas the other strains have a tendency to display a stronger signal from the smaller 650 bp band. This may suggest the smaller common band, which has been selected for use as a probe has had an insert into it at some point. This insert results in the appearance of the larger band of 720 bp size in some strains. From a comparison with the Cluster Analysis performed in Chapter 3, all those strains containing the 720 bp band appear in the larger *R. meliloti* cluster formed at the 80-100% level of similarity. This result appears to support the earlier findings that the *R. meliloti* species may

comprise two sub-groups. The 720 bp band may therefore be of use as a marker for differentiating amongst strains from both sub-groups.

Probe M1B:

Following hybridisation of Probe M1B back to RAPD profiles, the patterns shown in Figure 5.3 were obtained. No hybridisation with *R. leguminosarum* bv *trifolii* or *R. leguminosarum* bv *viciae* was observed with this probe. From Figure 5.3 it can be seen that a band of size 540 bp is present in *R. leguminosarum* bv *phaseoli* isolates *Rp*3603, *Rp*3604, *Rp*3606, *Rp*3607, *Rp*3609, *Rp*3613, *Rp*3614, *Rp*3615, *Rp*3617, *Rp*3620, *Rp*3622, and *Rp*3624. This band displays a strong signal in isolates *Rp*3603, *Rp*3604, *Rp*3606, *Rp*3607, *Rp*3609, *Rp*3617, *Rp*3620, and *Rp*3622. This 540 bp band is also the same size as the smallest band seen in strains of *R. leguminosarum* bv *phaseoli* when probed with M1A. A strong signal was seen in a larger band of size 630 bp in isolate *Rp*3618 and a weaker signal from a band of this size in isolates *Rp*3608 and *Rp*3619. The size and presence of the 630 bp band is identical to that observed when filters were hybridised with M1A. This indicates a certain degree of homology exists between the two regions of DNA forming M1A and M1B. These results again imply some DNA homology exists between strains from *R. meliloti* and *R. leguminosarum* bv *phaseoli*. However the level of hybridisation seen with M1B indicates a closer relationship exists than is suggested by results from probing with M1A. This cross-hybridisation between the species implies that M1B may not be of use as a species-specific marker.

R. meliloti M1B, when back-probed to *R. meliloti* reveals the presence of between one and three major bands, of size 720 bp, 650 bp and 595 bp, in the isolates. The smallest band being of equivalent size to M1B. The most common of these bands is the smallest (595 bp). This band is visible in fourteen isolates (it being unobservable in *Rm*2004 and absent from *Rm*2015). It is also assumed that the band is present in isolate *Rm*2000 (although not visible in Figure 5.3) as the probe was originally derived from this strain, thus making it present in fifteen of the seventeen isolates of *R.*

meliloti. The next smallest band (650 bp) is present in nine of the isolates, *Rm*2000, *Rm*2003, *Rm*2005, *Rm*2006, *Rm*2007, *Rm*2008, *Rm*2010, *Rm*2013, *Rm*2016. This size band is equivalent to that forming M1A which supports the suggestion that a degree of DNA homology exists between these probes.

The largest of the common bands is also present in nine isolates, *Rm*2000, *Rm*2003, *Rm*2005, *Rm*2007, *Rm*2008, *Rm*2010, *Rm*2011, *Rm*2013, *Rm*2016. This band is the predominant band in isolate *Rm*2003.

Four other bands appear in various strains of *R. meliloti*. Two isolates, *Rm*2005 and *Rm*2008, display the presence of a band of similar size to that of probe M1C (400 bp). Isolate *Rm*2002 displays the presence of a very small band of size 270 bp which is not seen in any other isolate of this species. The small size and rarity of this band may indicate that it was derived from the larger, common 595 bp band which has lost a 325 bp segment of DNA. This new 270 bp band is now presumably carried as an extraneous piece of DNA by the isolate *Rm*2002 only. The final two bands are of size 685 bp and 450 bp. These can be observed in strains *Rm*2015 and *Rm*2016 respectively.

This banding pattern indicates that M1B is common to most *R. meliloti* strains and some strains of *R. leguminosarum* bv *phaseoli*. The results also suggest that M1B shares some homologous DNA with M1A.

Probe M1C:

This probe, like the other two *R. meliloti* probes, M1A and M1B, showed no hybridisation with either *R. leguminosarum* bv *trifolii* or *R. leguminosarum* bv *viciae* isolates. The results of hybridising M1C to the *R. leguminosarum* bv *phaseoli* and *R. meliloti* RAPD profiles are shown in Figure 5.4. Hybridisation with *R. leguminosarum* bv *phaseoli* revealed a single band of size 630 bp in isolate *Rp*3618 only. When *R. meliloti* filters are probed a band of size 400 bp, which is equivalent to M1C, can be observed in eleven of the isolates. Isolate *Rm*2015 does not contain this band, however, it does reveal the presence of a larger band of size 685 bp instead which may

be the result of the insertion of a 285 bp piece of DNA into the 400 bp band. Isolates *Rm*2000, *Rm*2003, *Rm*2005, *Rm*2008, *Rm*2011 and *Rm*2013 also display the presence of a larger band of 720 bp. This sized band is also observed when hybridisations with both M1A and M1B are made. Its appearance with all three probes suggests that M1A and M1B share some DNA homology with one part of this band whilst M1C is homologous with a different part. This is supported by M1A and M1B sharing other homologous bands, while showing little or no homology with M1C. There is no evidence of a signal from the M1A and M1B sized DNA bands. This, again, indicates that there is no homology between M1C and M1A and M1B.

Probes V1D and T1E:

When probes V1D and T1E were hybridised back to RAPD profiles the results shown in Figure 5.5 were obtained. Both these probes, although varying in size by about 40 bp (530 bp [V1D] and 490 bp [T1E]), hybridise in an identical manner thus strongly supporting the suspicions, based on RAPD profile observations, that they are from identical parts of the genome from the two respective biovars. This supports earlier findings (Chapter 3, Dooley *et al.*, 1993) that it is not possible to differentiate between isolates from these two biovars using RAPD profile analysis. No hybridisation was observed with the species *R. meliloti* or with the *R. leguminosarum* bv *phaseoli* isolates. The lack of hybridisation, from either probe, to the *R. leguminosarum* bv *phaseoli* strains lends further support to the theory that this biovar may not be as closely related to the biovars *trifolii* and *viciae* as they are to each other. Similar reports have been made previously (Chapter 3; Segovia *et al.*, 1991; Dooley *et al.*, 1993). From Figure 5.5 it can be seen that bands of 530 bp and 490 bp are revealed in the majority of the isolates of *R. leguminosarum* bv *viciae* and *R. leguminosarum* bv *trifolii* respectively. These are identical in size to DNA used to produce the probes V1D (530 bp) and T1E (490 bp). This piece of DNA is present in all isolates (except *R. leguminosarum* bv *viciae* *Rv*1015 and *Rv*1018 and *R. leguminosarum* bv *trifolii* *Rt*3, *Rt*35, *Rt*7D5 and *Rt*R1.3) from these two biovars. The signal intensity from the

strains varies greatly and may be the result of a varying copy number of the DNA band in each isolate, which suggests a plasmid borne piece of DNA.

***Conclusions*:**

Based on RAPD probing the following general observations can be drawn. (i) The two probes M1A and M1B appear to share a portion of homologous DNA which they do not share with M1C. (ii) The two *R. leguminosarum* probes, V1D and T1E, appear to be identical except for a small (40 bp) size difference. (iii) These two probes also appear to be highly species-specific with no cross-hybridisation with non *R. leguminosarum* strains. They also do not hybridise with the isolates from *R. leguminosarum* bv *phaseoli* which tends to infer that this biovar is not as closely related to the other two biovars as they are to each other. Reports to this effect have been made previously (Chapter 3; Segovia *et al.*, 1991; Dooley *et al.*, 1993). (iv) None of the *R. meliloti* probes display cross-hybridisation to isolates of *R. leguminosarum* biovars *trifolii* and *viciae* although they do show a limited amount of cross-reactivity with a few strains of *R. leguminosarum* bv *phaseoli* This again supports the suggestion that *R. leguminosarum* bv *phaseoli* may not be as closely related to *R. leguminosarum* biovars *trifolii* and *viciae* as they are to each other. Cross-reactivity such as this also indicates a degree of homology between strains of these two species. This may be possible as there have been reports of DNA exchange between isolates of these two species (Djordjevic *et al.*, 1983; Broughton *et al.*, 1987). (v) The observation that not all the strains of *R. leguminosarum* bv *phaseoli* hybridise with the *R. meliloti*-derived probes tends to indicate that this group is not homogenous in its make up. This has been suggested by others (Chapter 3; Pinero *et al.*, 1988; Dooley *et al.*, 1993) (vi) M1B displays the greatest amount of cross-reactivity with isolates of *R. leguminosarum* bv *phaseoli* which indicates that this probe may be of little use as a species-specific probe. (vii) The overall levels of species-specific hybridisation displayed by the probes in this study appears to be high. This may not be the case, however, as the RAPD profile is a construct comprising a small, select part of the

bacterial genome thus identical genes from a second species may not be amplified owing to small differences in the primer binding sites.

As probing investigations have so far concentrated on analysis of Southern blots of amplified DNA, which may not be a total representation of a strains genomic DNA content, it would be interesting to examine total bacterial genomic DNA using these probes. This could be achieved by probing colony blots or restriction digests of total genomic DNA.

Tables and Figures

Table 5.1: This shows details of the species-specific probes used in this study. Specific species, derivative strains, probe sizes and relative frequency of the RAPD band within each species are shown.

Probe	Species	Derivative strain	Size (bp)	Frequency in population[1]
M1A	*R. meliloti*	2000	650	71 %
M1B	*R. meliloti*	2000	600	59 %
M1C	*R. meliloti*	2000	400	77 %
V1D	*R. leg.* bv *viciae*	1001	530	83 %
T1E	*R. leg.* bv *trifolii*	JJD4	490	78 %

[1] Frequency calculated from the presence of the band in isolates following DNA amplification.

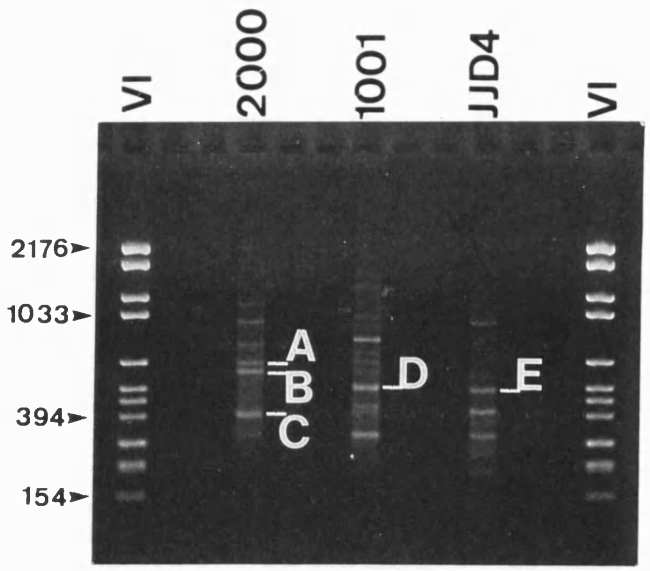


Figure 5.1: RAPD fingerprints of the three strains used for the isolation of species-specific probes are shown. The common bands selected for constructing the probes, M1A (A), M1B (B), M1C (C), V1D (D) and T1E (E) are arrowed. Marker sizes (Marker VI, B. Mannheim) (bp) are shown.

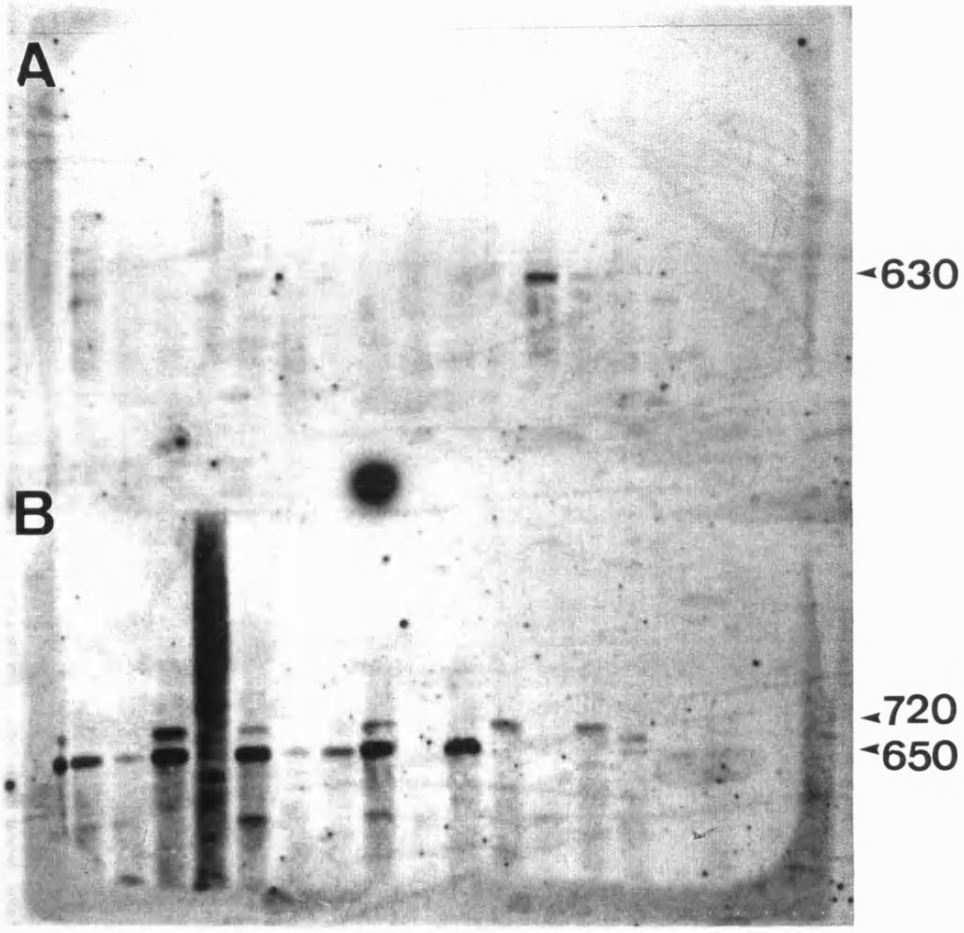


Figure 5.2: Results of hybridising RAPD profiles from isolates of *R. leguminosarum* bv *phaseoli* (A) and *R. meliloti* (B) with probe M1A. Strains were loaded in order, as shown in Table 2.1 (page 46). Size markers (Marker VI, B. Mannheim) are in bp.

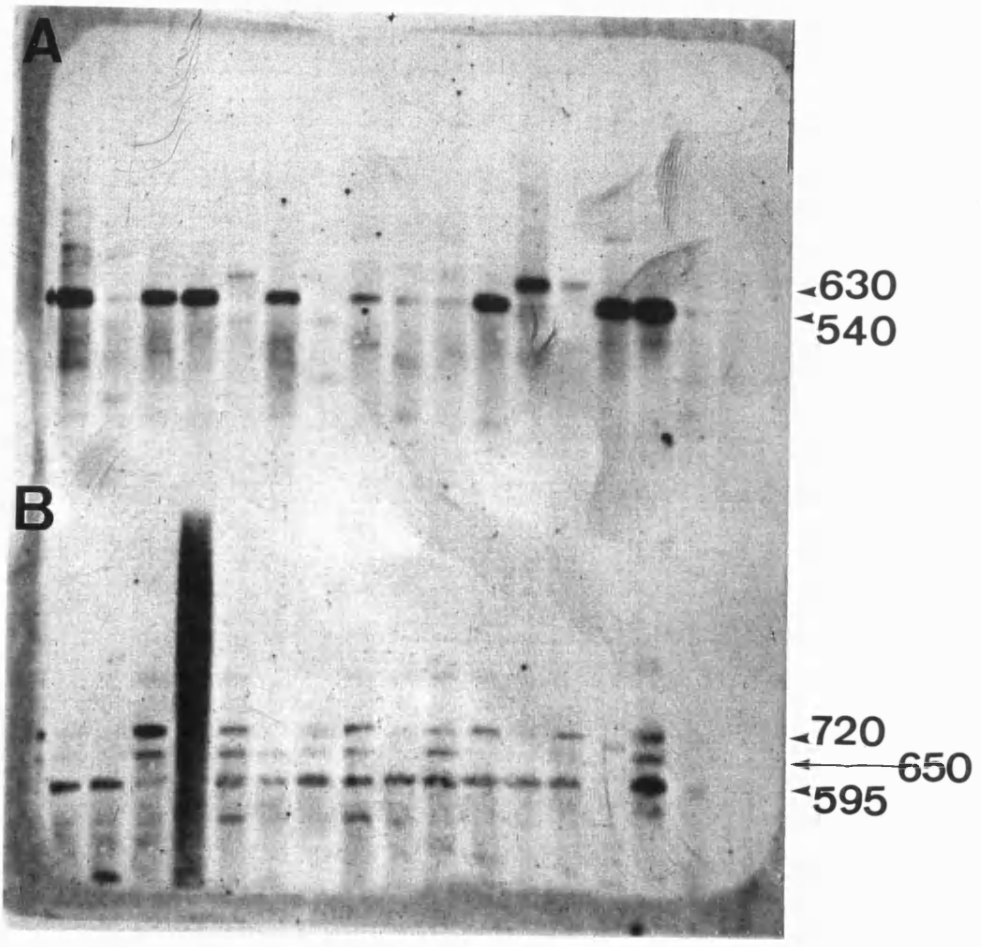


Figure 5.3: Results of hybridising RAPD profiles from isolates of *R. leguminosarum* bv *phaseoli* (A) and *R. meliloti* (B) with probe M1B. Strains were loaded in order, as shown in Table 2.1 (page 46). Size markers (Marker VI, B. Mannheim) are in bp.

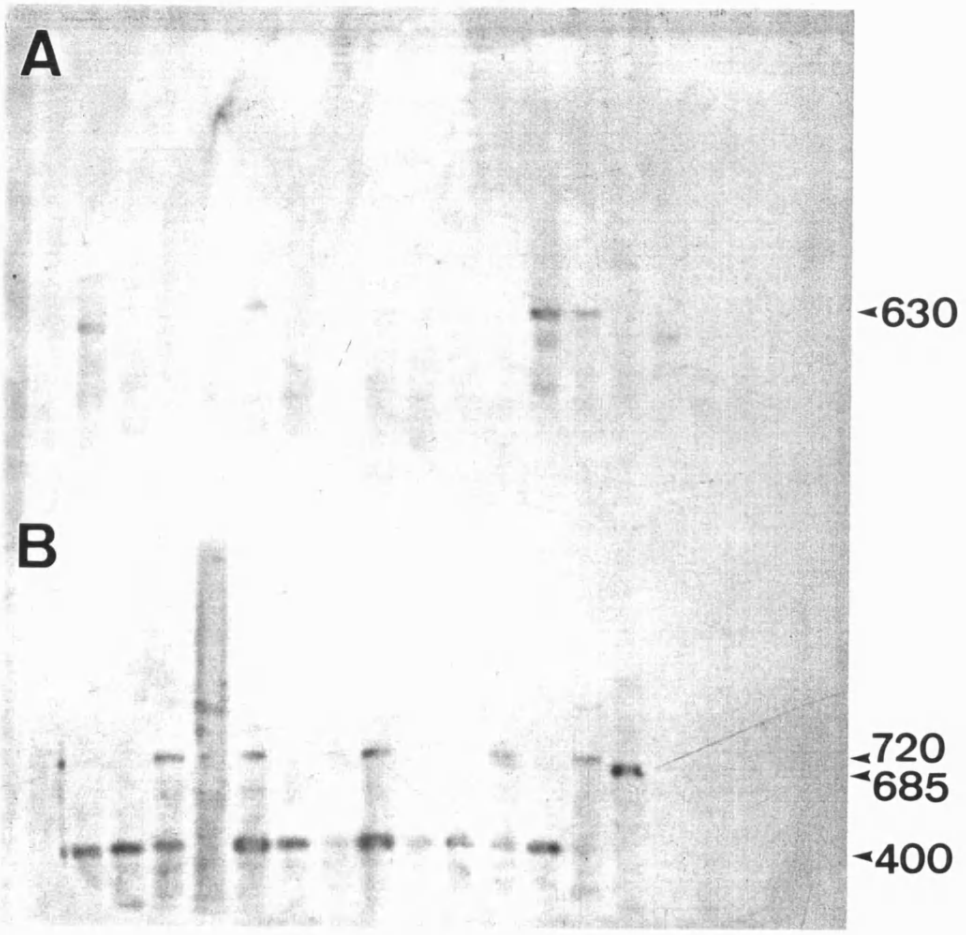


Figure 5.4: Results of hybridising RAPD profiles from isolates of *R. leguminosarum* bv *phaseoli* (A) and *R. meliloti* (B) with probe M1C. Strains were loaded in order, as shown in Table 2.1 (page 46). Size markers (Marker VI, B. Mannheim) are in bp.

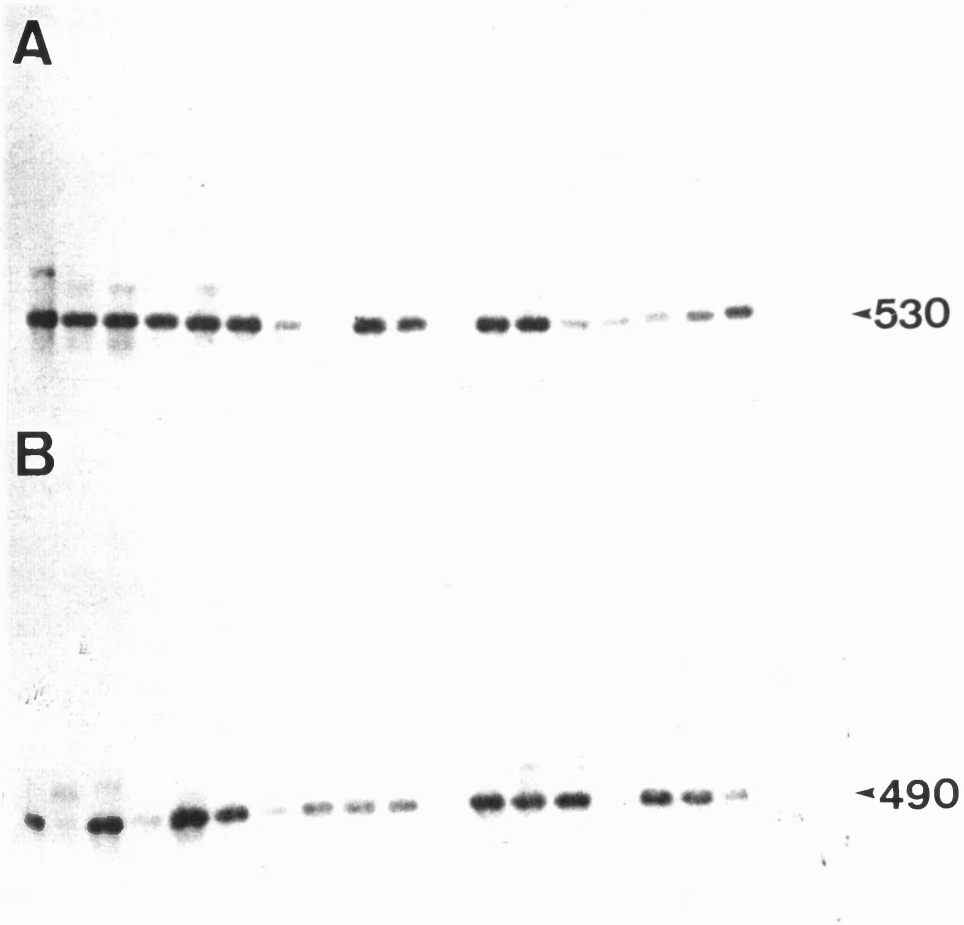


Figure 5.5: Results of hybridising RAPD profiles from isolates of *R. leguminosarum* bv *viciae* (A) and *R. leguminosarum* bv *trifolii* (B) with either of the probes V1D or T1E. Both probes reacted in a similar manner although isolated from different biovars. Strains were loaded in order, as shown in Table 2.1 (page 46). Size markers (Marker VI, B. Mannheim) are in bp.

Further Assessment of the Specificity of RAPD-Derived Probes by Hybridisation to Colony Blots of *Rhizobium* Isolates.

Material from this Chapter has been included in the paper:

DOOLEY J.J., HARRISON S.P., MYTTON L.R., DYE M., CRESSWELL A., SKØT L. & BEECHING J.R. The Use of Colony Blots to Rapidly Assess the Value of Potential RADP-Derived Strain-Specific Rhizobium Probes. Submitted to FEMS Microbiology Letters.

Abstract

The potential species-specific probes M1A, M1B, M1C, V1D and T1E, which have been assessed previously by hybridisation to Southern blots of RAPD profiles, were hybridised against total genomic DNA, from the species *Rhizobium leguminosarum*, *R. meliloti* and *Bradyrhizobium*, by the use of colony blotting. The *R. meliloti* probe, M1B, although displaying cross-reactivity with the RAPD blots of *R. leguminosarum* bv *phaseoli*, was screened in order to assess its distribution within the *Rhizobium* and *Bradyrhizobium* species in general. The results observed indicate that only one species-specific probe, M1C, showed potential for use in this area and then only in a limited capacity.

Colony blots were also used to assess the value of a series of potentially strain-specific probes which had been isolated from agarose containing RAPD amplifications of *Rhizobium* DNA. These probes had been selected on the basis of their appearance in only one isolate from the 84 isolates employed in this study. Following colony hybridisation only one strain-specific probe, T37-3, showed a high potential for this type of use although results obtained from analysis of T37-2 indicate that this probe maybe useful in certain situations. Analysis of four additional strain-specific probes (T1-2, T1-3, V1-1 and T37-1) indicated they although they were of no use as strain-specific probes they were potentially of use as species-specific probes.

Introduction

The use of colony blotting as a method to assess the value of DNA probes has been reported by Manulis, (1992) who evaluated a DNA probe, for the detection of *Erwinia herbicola*, by comparing colony blot results with enzyme-linked immunosorbent assay (ELISA) and pathogenicity test results. Manulis found that colony hybridisation was specific for gall forming pathogens and sensitive enough to detect 100 colony forming units (CFUs) after enrichment. Daire *et al.*, (1992) used colony blotting to screen a cloned DNA library for the presence of an insert from a non-cultivable mycoplasma-like organism (MLO). They then used the clones as probes for detecting MLO in field samples of grapevine. Colony blotting has also been employed for the identification of *Rhizobium* species. Hodgson and Roberts, (1983) used colony hybridisation to identify *R. leguminosarum* bv *trifolii* isolates obtained from nodules of subterranean clover. They used total genomic DNA from individual strains as probes to determine nodule occupancy. A similar method was employed by Cooper *et al.*, (1987) to determine if nodules from *Lotus pedunculatus* were occupied by *R. loti* or *Bradyrhizobium* species. They found that total genomic probes were not strain-specific, revealing varying degrees of cross-hybridisation with isolates of the same genus. However, they also noted that there was no cross-reactivity between the two genera, *Rhizobium* and *Bradyrhizobium*. Fredrickson *et al.* (1988) also used this technique, in conjunction with several others, for the enumeration of *Rhizobium* and *Pseudomonas* populations in soil. However, this group used Tn5 as a probe to detect the isolates. All these methods have employed radioactive labelled DNA probes. The use of non-isotopic DNA probes with colony blotting has been reported by Lonvaud-Funel *et al.*, (1991) who employed Dig-labelled total genomic DNA probes to study the evolution of lactic acid bacterial species in grape musts and wines. The use of Dig-labelled, RAPD-derived, species-specific probes to screen RAPD profiles is described in Chapter 5. The present study hopes to confirm the specificity of these species-

specific probes and to assess the value of a series of RAPD products, observed in single isolates only, as strain-specific *Rhizobium* probes.

Materials and Methods

Probe Construction:

Potential *Rhizobium* strain-specific probes were identified by a close examination of RAPD fingerprints produced using the primers SPH1 (Chapter 3) and SPH3 and 7 (Chapter 4). No strain-specific bands were observed with the primer SPH3 alone. Bands appearing in only one isolate from each species or biovar were selected for screening to ascertain their use as strain-specific probes. This method was performed prior to the publication of the technique of combined subtraction-hybridisation and DNA amplification for strain-specific DNA isolation (Bjourson and Cooper, 1992).

Bands of interest were extracted from 1.5% agarose gels and Dig-labelled using the band-stab technique described in Chapter 2. These probes are described in Table 6.1.

Colony blotting:

Fixation of DNA to nylon membranes:

Discs (85 mm diameter) of Zeta-probe GT, nylon membrane (Biorad) were cut from the sheet and sandwiched between two sheets of filter paper (Whatman 3MM) prior to sterilisation by autoclaving. The discs were laid, avoiding air bubbles, onto TY agar plates before the *Rhizobium* and *Bradyrhizobium* strains were plated out by picking small inocula from an overnight culture grown on TY agar plates and placing them in a grid pattern on the Zeta-probe (Figures 6.1a and 6.1b). When probing with species-specific probes all 84 isolates were employed. These isolates were laid out on the Zeta-probe in a grid pattern as shown in Figure 6.1a. Each species was placed onto a separate filter. When strain-specific probes were screened by colony blotting only those strains from which probe DNA was isolated were used. The layout of these

strains is shown in Figure 6.1b. The plates were incubated at 27° C until colonies were 2-5 mm diameter before cells were lysed and DNA fixed to the filters using the following procedure. The membranes were placed, colony side up, onto filter paper soaked in 10% w/v SDS for 5 min. They were transferred to filters soaked in lysis solution (1.5 M NaCl, 0.5 M NaOH) for 5 min before being transferred to neutralising solution (1.5 M NaCl, 0.5 M Tris-Cl, pH 8.0) for 5 min. Finally, the membranes were transferred to filters soaked in 2x SSC for 2x 5 min, then dried for 30 min on 3 MM filter paper. DNA was fixed to the membranes by baking at 80° C for 2 hrs.

Hybridisation:

Filters were hybridised with the various probes as described in chapter 2, but a greater level of stringency was used in the final washes (50° C and 0.1x SSC). Those probes displaying a high level of strain specificity were further screened against a larger selection of isolates from their own species or biovar.

Results and Discussion

Species-specific probes:

The results obtained from probing colony blots with the various species-specific probes are shown in Figures 6.2 to 6.5. The filters carry DNA from the species *R. leguminosarum* bv *viciae* (V), *R. leguminosarum* bv *trifolii* (T), *R. leguminosarum* bv *phaseoli* (P), *R. meliloti* (M), *Bradyrhizobium* (B) and non-*Rhizobium* (nR) respectively. There are also three control strains at the bottom of each filter comprising *R. leguminosarum* bv *trifolii* isolate *Rt*JJD4 (A), *R. leguminosarum* bv *viciae* isolate *Rv*1001 (B) and *R. meliloti* isolate *Rm*2000 (C). The probes T1E, V1D and M1A, M1B and M1C were produced by Dig-labelling the common bands found in RAPD profiles from these three strains. The control sample D is eukaryotic DNA (herring sperm DNA). In all hybridisations no positive result was obtained with the herring sperm DNA.

Probe M1A:

Figure 6.2 shows the results of probing with M1A which was isolated from *R. meliloti*, strain *Rm*2000. From an examination of the results it can be seen that this probe appears to be fairly species-specific displaying little cross-hybridisation with non *R. meliloti* isolates, whilst maintaining a high level of hybridisation with isolates of the species *R. meliloti*. There is, however, a strong hybridisation to four of the *Bradyrhizobium* isolates suggesting a degree of conservation exists between these two unrelated species. This is in contrast to the results of back-probing to RAPD fingerprints (Chapter 5) where none of the species-specific probes hybridised to any isolates of the genus *Bradyrhizobium*. When colony blot and RAPD profile analysis (Chapter 3, Dooley *et al.*, 1993) results are compared it can be seen that four of the five *Bradyrhizobium* isolates linking to the *R. meliloti* cluster at the 75-80 % level of similarity are the same as those now hybridising to the *R. meliloti* species-specific probe. This observation, therefore, lends support to the earlier RAPD findings

(Chapter 3, Dooley *et al.*, 1993) which indicate a degree of homology exists between one of the sub-groups forming the species *Bradyrhizobium* and the species *R. meliloti*. The four strains showing positive results with the M1A are not of the recognised species *B. japonicum*. Three of the strains were isolated from lupins and one from a *Cicer* species. The geographic origins of the strains are also quite diverse, one strain having been isolated from the U.S.A, one from the U.K. and one from New Zealand. The remaining strain is of unknown origin. Reports that the species *Bradyrhizobium* may be an aggregation of two or more strain types have been made by several groups, Stanley *et al.*, (1985) and Kuykendall *et al.*, (1988). A report by Sadowsky *et al.*, (1987a) suggests that *R. fredii* (now reclassified as *Sinorhizobium fredii* [Chen *et al.*, 1988]) may represent an evolutionary link between *Bradyrhizobium* and *R. meliloti*. It may be that the four strains classified here as *Bradyrhizobium* are in fact isolates of the genus, *S. fredii*. This would explain the observations seen here but without further details about the strains in question it is not possible to confirm this.

M1A displays no hybridisation with the non-*Rhizobium* (TG1-TG14) isolates and low levels of hybridisation with some *R. leguminosarum* isolates (as compared to the signal from the control strain). Only one isolate of *R. leguminosarum* bv *trifolii* (*Rt*JJD4), two strains of *R. leguminosarum* bv *viciae* (*Rv*1004 and *Rv*1021) and one *R. leguminosarum* bv *phaseoli* isolate (*Rp*3613) show any sign of hybridisation with this probe, which suggests a degree of DNA conservation exists between these two species. This is in contrast with results of Southern blots of RAPD profiles where *R. leguminosarum* bv *phaseoli* isolates *Rp*3604, *Rp*3607, *Rp*3608, *Rp*3619 and *Rp*3622 hybridised weakly to M1A. This variation in hybridisation patterns could be accounted for in the stringency wash temperatures, which are considerably higher with the colony blots. Any weakly binding probes observed on the RAPD blots would be removed in this case. Results where strains display hybridisation signals on the colony blots but not on the RAPD blots indicate that part of the genome, which is homologous to the probe, is not being amplified by the primer SPH1 in these strains. This highlights the

obvious differences between those strains which possess or lack the SPH1 primer binding site.

Probe M1B:

Results of probing colony blots with M1B are shown in Figure 6.3. This probe hybridises strongly with all isolates from both the genera, *Rhizobium* and *Bradyrhizobium*, and also displays a strong signal with five of the ten non-*Rhizobium* isolates used in this study. This strong hybridisation to everything, except the herring sperm DNA, intimates that the 595 bp band amplified during the RAPD reaction is from a part of the genome found in the majority of soil-borne bacteria. When M1B was hybridised to RAPD profiles (Chapter 5) no annealment with strains from the species *Bradyrhizobium* or *R. leguminosarum* biovars *viciae* and *trifolii* was observed. However, the high number of *R. leguminosarum* bv *phaseoli* strains which did hybridise to this probe insinuated that the probe may not be totally species-specific. The results from colony blotting, which reveal the probe to hybridise to nearly all the *Rhizobium* and *Bradyrhizobium* strains, support this conclusion. Hybridisation, by this probe, to the non *Rhizobium* isolates also render the probe useless as a general *Rhizobium* probe.

Probe M1C:

The results of hybridising colony blots with M1C are shown in Figure 6.4. From an examination of Figure 6.4 it can be seen that the *R. meliloti* strains hybridise well with M1C, although a weak signal is obtained from isolates *Rm*2003, *Rm*2006, *Rm*2015 and *Rm*2017 and practically no hybridisation can be observed with strain *Rm*2016. It can also be seen that the four *Bradyrhizobium* isolates which hybridised to M1A also hybridise with M1C. The remaining filters reveal no or very weak hybridisation signals (as compared to the control strains) from a few isolates only. Those strains revealing possible homology with M1C are *R. leguminosarum* bv *phaseoli* isolates *Rp*3613, *Rp*3614 and *Rp*3615 and the non *Rhizobium* strains TG3

and TG4. When RAPD blots were hybridised with this probe only isolate *Rp*3618 revealed any homology to the probe. This was in the form of a single 630 bp band which was, however, of comparable signal intensity to those observed in the *R. meliloti* strains. These results are therefore similar to findings made when RAPD profiles were hybridised with this probe, in that only *R. leguminosarum* bv *phaseoli* isolates indicate any homology to this probe. This may be due to the mixed strain types reported to compose *R. leguminosarum* bv *phaseoli* (Chapter 3; Pinero *et al.*, 1988; Dooley *et al.*, 1993). It may also be that these three strains are of the recently proposed species, *R. tropici* (Martínez-Romero *et al.*, 1991). However, the colony blot results tend to suggest that M1C is not as species-specific as was inferred from results of RAPD blots. This probe is therefore probably of little use for the identification of *R. meliloti* strains soil samples containing a diverse population of bacterial strains. It may, however, be of use in studies in a controlled environment where only a limited number of bacterial strains are under investigation.

Probes V1D and T1E:

The probes V1D and T1E hybridise in an identical manner on the colony blots and so have been dealt with together. The high degree of similarity between these two probes, as observed from both RAPD profiles and colony blots, strongly indicates that, although varying in size by 40 bp, they are from identical parts of the respective genomes. The identical hybridisation patterns observed with both probes on colony blots reiterates conclusions made from the RAPD profile analysis (Chapter 3) that these two biovars are closely related. However, the evidence of a slight size difference between the probes does indicate that variations exist between the two biovars.

The results of probing colony blots with either probe are shown in Figure 6.5. As can be seen from Figure 6.5, all the isolates of *R. leguminosarum* biovars *trifolii* and *viciae* hybridise with the probes. This, again, highlights the great similarity amongst strains of these two biovars. Results of hybridisation to *R. leguminosarum* bv *phaseoli* strains reveals that twelve of the isolates (*Rp*3603, *Rp*3604, *Rp*3605, *Rp*3607,

*Rp*3608, *Rp*3609, *Rp*3611, *Rp*3613, *Rp*3617, *Rp*3620, *Rp*3624 and *Rp*3626) hybridise to the probes to some degree. This result, which is contrary to that from RAPD blots, where no hybridisation to *R. leguminosarum* bv *phaseoli* isolates occurred, indicates the existence of some DNA homology between the strains of these three biovars. Such an observation would be expected between biovars of the same species, however, the extent of hybridisation suggests that the isolates comprising *R. leguminosarum* bv *phaseoli* are not as closely related to the other biovars (*trifolii* and *viciae*) as these biovars are to each other. This is similar to previous reports to this effect (Chapters 3 and 4; Segovia *et al.*, 1991; Dooley *et al.*, 1993). Hybridisation with isolate *Rp*3613, which also hybridised to M1C, suggests that isolate may be an evolutionary link between the two species *R. leguminosarum* and *R. meliloti*.

Results obtained from hybridising the *R. meliloti* filter with these two probes reveals that at least one strain (isolate *Rm*2005) shares a degree of DNA with the species *R. leguminosarum*. It is possible that this shared DNA was passed by genetic exchange which has been reported to occur between strains of these species. Djordjevic *et al.*, (1983) have reported that plasmid exchange can be induced between isolates of *R. meliloti* and *R. leguminosarum* under laboratory conditions. The exchange of genetic material in the rhizosphere of the recipient strains host plant has been recorded by Broughton *et al.*, (1987). There have not, however, been reports of exchange between the species *R. leguminosarum* and the genus *Bradyrhizobium* which could account for the positive result observed when the *Bradyrhizobium* isolates *Br*3205, *Br*3212, *Br*3213, *Br*3413 and *Br*3828 were probed with V1D and T1E. These five strains are from diverse geographic origins including Brazil, England, USA and New Zealand and only isolate *Br*3413 is of the recognised species *B. japonicum*. It may be that the four strains (*Br*3205, *Br*3212, *Br*3213 and *Br*3828) are from the species *S. fredii* (Chen *et al.*, 1988) which has been suggested as a link between *Bradyrhizobium* and *R. meliloti* (Sadowsky *et al.*, 1987a). The results presented here may indicate an evolutionary link between *R. leguminosarum* and *Bradyrhizobium* or the new genus *Sinorhizobium*.

The two probes, V1D and T1E, also hybridise with two non *Rhizobium* strains, TG4 and TG6, which suggests that these strains share some homology with the species *R. leguminosarum*. This final piece of evidence reveals that these two probes are not as species-specific as previous results have implied.

Overall observations from the results of hybridisation of probes M1A, M1B, M1C, V1D and T1E to colony blots of the 84 *Rhizobium* and *Bradyrhizobium* strains and the ten non-*Rhizobium* isolates used here indicate that (i) back-probing to RAPD profiles can produce a distorted view of the DNA homology which exists amongst strains of diverse species. This is due to the nature of the RAPD reaction which only considers select parts of the whole genome. Those isolates which do not contain the primer sites will, therefore, not be considered in RAPD reactions or in any subsequent analysis even though they may carry the intact gene. (ii) The *R. meliloti* probe, M1A, appears to hybridise strongly with all the *R. meliloti* isolates which denotes a strong homology amongst the strains of this species even though findings by others (Young *et al.*, 1985; Eardly *et al.*, 1990; Dooley *et al.*, 1993) have indicated that this species may comprise two sub-groups. The evidence here, therefore, implies that even if two sub-groups exist within *R. meliloti* a certain amount of DNA is shared between them. (iii) The evidence that cross-reactivity between strains of the species *R. leguminosarum* and *R. meliloti* occurs implies a degree of homology exists between these two *Rhizobium* species. Evidence of plasmid transfer between isolates of both species has been reported previously (Djordjevic *et al.*, 1983; Broughton *et al.*, 1987). This could account for the observations made here. (iv) Evidence obtained from RAPD profile analysis (Chapter 3) and RAPD blotting (Chapter 5) that a link between *Bradyrhizobium* and *R. meliloti* exists is further confirmed by results obtained from colony blotting. (v) The colony blotting results described here also indicate a possible evolutionary link between the species *R. leguminosarum* and the genus *Bradyrhizobium*. (vi) The highly similar hybridisation patterns observed with probes T1E and V1D further supports the belief that they are from highly conserved areas of the genome found in their respective biovars, *R. leguminosarum* biovars *trifolii* and

viciae. (vii) However, the results from colony blotting counteract the previously made suggestion (based on RAPD blots) that the probes V1D and T1E may be of use for species-specific studies. (viii) Finally, results described here point to the use of M1C for species-specific work in a controlled environment as it displays a limited amount of cross-reactivity to other species.

Strain-specific probes:

From a close examination of the RAPD profiles obtained using the various primers (Chapters 3 and 4) a selection of potential strain-specific probes were isolated and Dig-labelled using the band-stab method of Bjourson and Cooper (1992). These bands and their respective derivative strains are shown in Table 6.1. No strain-specific probes were identified from SPH3 primed RAPD reactions.

In order to identify strain-specific probes it is desirable to employ stringency wash conditions which result in few isolates displaying positive signals following hybridisation with the probe. To achieve this, very stringent conditions (50° C, 0.1x SSC) for the post-hybridisation washes were found to be optimal as this reduced the number of fainter signals observed when lower stringency washes were employed. The results obtained following washing at this high level of stringency are shown in Figures 6.6 to 6.8.

Probes derived from SPH1 primed RAPD reactions:

Figure 6.6 shows the results obtained when a series of probes derived from RAPDs produced with primer SPH1 were screened. The filters A-D have been probed with the potential *R. leguminosarum* bv *trifolii* strain-specific probes T1-1, T1-2, T1-3 and T1-4 respectively. An overall examination of the filters reveals a variation in hybridisation patterns between these probes. However, in general, none of them show great potential as strain-specific probes. It was also noted that none of the probes T1-

1, T1-2, T1-3 or T1-4 reveal evidence of homology with isolates of the species *R. meliloti*.

Probe T1-1 hybridises strongly with isolates *Rv*1001, *Rv*1023 and *Rp*3604 but slightly less well to isolates *Rp*3603, *Rp*3608, TG1 and TG3. It, however, fails to display such levels of hybridisation with its own derivative strain, *Rt*JJD4. This may be the result of a poor fixing of DNA, from the strain *Rt*JJD4, to the filter. This observation may also be the result of small genomic variations between these isolates which has resulted in only isolate *Rt*JJD4 displaying the presence of the T1-1 sized band. The high number of strains showing positive hybridisation results with this probe, whilst failing to reveal its presence after DNA amplification, suggests that isolate *Rt*JJD4 contains a mutation which has resulted in the formation of a primer site. The low level of strain-specificity observed during hybridisation tests with this probe mean it is of little use for studies of *Rhizobium*. Probe T1-2 hybridises well with *Rt*WPBS3 (its derivative strain) but also shows good hybridisation to three other *R. leguminosarum* bv *trifolii* isolates (*Rt*JJD4, *Rt*JJD15 and *Rt*162S7A). This probe also hybridised to three isolates of *R. leguminosarum* bv *viciae* (*Rv*1001, *Rv*1007 and *Rv*1023) and two strains of *R. leguminosarum* bv *phaseoli* (*Rp*3604 and *Rp*3608). The inconsistency in probing patterns observed with this probe render it of no use as a strain-specific probe although it may be of use as a species-specific probe. Probe T1-3 displays a good level of hybridisation with its derivative strain, *Rt*WPBS3, but reveals a stronger hybridisation signal with strain *Rt*162S7A. This again may be the result of a problem with DNA fixation to the filter or has been caused by an altered primer binding site. From the results of RAPD analysis (Chapter 3) it was found that these two isolates fall within the same region of the PCO plot (Figure 3.2) and are grouped together at the 75-80% level of similarity. T1-3 also hybridises with isolates *Rt*JJD15, *Rv*1001, *Rv*1023, *Rp*3604 and *Rp*3608, all of which are strains of the species *R. leguminosarum*. This indicates that a degree of DNA homology exists between certain strains of the three biovars of *R. leguminosarum* and implies that the T1-3 could be useful for species-specific identification. The final probe from this series, T1-4,

hybridises well with strains *Rt*162S7A (its derivative) and *Rv*1023. There is also evidence of hybridisation with strains *Rt*JJD15, *Rv*1001, *Rp*3603, *Rp*3604 and the non-*Rhizobium* strain TG3. Hybridisation with the non-*Rhizobium* strain implies that this probe is of no use as either a strain-specific or a species-specific probe. It may, however, indicate that the unclassified strain TG3 shares a degree of homologous DNA with the species *R. leguminosarum*.

Figure 6.7 shows the results of colony blotting with potential strain-specific probes of the species *R. leguminosarum* bv *viciae* (V1-1 and V1-2) and *R. meliloti* (M1-1). The filters A and B have been probed with the *R. leguminosarum* bv *viciae* probes, V1-1 and V1-2, respectively. Filter C has been hybridised with the *R. meliloti* probe, M1-1.

The probe V1-1 appears to be of no use as either a species-specific or strain-specific probe as it hybridises well with all isolates except *Rt*JJD4, *Rt*WPBS3, *Rt*46rif, *Rv*1014 and TG10. Isolates to which the probe hybridises include those from the species *R. meliloti* (*Rm*2000, *Rm*2001 and *Rm*2004) and the non-*Rhizobium* strains (TG1 and TG3). Probe V1-1 appears, from the results of colony blotting, to be from a gene which is found in the majority of soil-borne bacteria. However, variations in genomic sequence have meant that it is only apparent in certain strains following DNA amplification with primer SPH1. Probe V1-2 also offers little hope of being useful as a strain-specific probe but may have a potential use as a species-specific probe. This probe displays strong hybridisation signals from isolates *Rv*1023 (its derivative strain), *Rt*JJD4, *Rt*JJD15, *Rp*3603 and *Rp*3604. There are also weaker signals from the remaining isolates of the species *R. leguminosarum* except isolates *Rt*46rif and *Rp*3619. No hybridisation with the *R. meliloti* or non-*Rhizobium* isolates was observed. These observation again support the present classification of these three biovars under the one species, *R. leguminosarum*, (Elkan, 1992). The results also imply that V1-2 is from a region of the *R. leguminosarum* genome which has no homology with the species *R. meliloti*.

The potential *R. meliloti* strain-specific probe, M1-1, shows a high level of cross-reactivity with other isolates from this species thus rendering it useless as a strain-specific probe. This probe also hybridises strongly with the *R. leguminosarum* bv *trifolii* isolate, *Rt*162S7A and to a lesser extent to the *R. leguminosarum* bv *viciae* isolate *Rv*1023. This cross-reactivity indicates a high level of conservation exists between these two strains (*Rt*162S7A and *Rv*1023) and the species *R. meliloti*. This could be the result of genetic exchange which has been reported to occur between isolates from these two species (Djordjevic *et al.*, 1983; Broughton *et al.*, 1987). A degree of DNA conservation also exists between the *R. meliloti* isolates and two of the non-*Rhizobium* isolates as indicated by hybridisation of this probe to the non-*Rhizobium* isolates TG1 and TG3.

Probes derived from SPH3 and SPH7 primed RAPD reactions:

Figure 6.8 shows the results of hybridisation with the potential strain-specific probes identified from RAPDs produced using the primers SPH3 and SPH7 in a dual primed reaction. The filters A to D have been hybridised with the probes T37-1, T37-2, T37-3 and T37-5 respectively. The probe T37-4 was not screened as it displayed a series of smaller, nested sub-bands when the original 2.2 Kb band was amplified.

The probe T37-1 displays a high level of hybridisation with its derivative strain, *Rt*162S7A, but also shows cross-reactivity with the strain *Rv*1023. The probe also exhibits lower levels of cross-hybridisational activity with several other isolates (*Rt*JJD15, *Rv*1001, *Rp*3603 and *Rp*3604) of the species *R. leguminosarum*. This cross-hybridisation renders the probe useless for strain-specific studies, although it may still be of use for species-specific work as there is no evidence of hybridisation with the *R. meliloti* isolates. The probe T37-2 hybridises strongly to its derivative strain, *Rt*JJD15, but also hybridises as efficiently to the *R. leguminosarum* bv *viciae* isolate, *Rv*1023. This would suggest a high level of DNA conservation amongst the two strains, *Rt*JJD15 and *Rv*1023, which is not unexpected between isolates from two closely

related biovars of the same species. However, this cross-reactivity again renders this probe of no use for strain-specific work.

Figure 6.8, filter C, shows the results of using the probe T37-3. It can be seen that the probe hybridises well with its derivative strain, JJD15, and shows no cross-hybridisational activity with any other strains from any species or biovar. Of all the bands screened for their potential use as strain-specific probes T37-3 appears to be the best. This strain and probe combination therefore indicate that it should be possible to use unmodified bacterial strains, and their specific probes, for studies of inoculum longevity and movement within natural environments. If a selection of probes could be identified for a single isolate it may also be possible to study genetic exchange between bacteria. However, work such as this would require a larger selection of probes to make it feasible. The method of subtraction-hybridisation and DNA amplification (Bjourson *et al.*, 1992) provides a means of efficiently producing a selection of probes for such studies. By combining the method of Bjourson *et al.*, (1992) with the RAPD reaction it should be possible to produce strain-specific probes, for a large selection of isolates, relatively quickly.

The final probe examined for its potential use as a strain-specific probe was T37-5. This probe displays a very poor hybridisation pattern at this high level of stringency washing. From an examination of Figure 6.8, filter D, it can be seen that the probe fails to display any signal from its derivative strain, *Rt*46rif. This, however, may be due to a low amount of DNA fixing to the filter as the strain grew poorly. However, it can also be seen that the probe produces weak hybridisation signals from isolates *Rv*1023, *Rm*2000, *Rm*2001 and *Rm*2009. The last three isolates are of the species *R. meliloti* which tends to imply a degree of homology exists between the two species. It also indicates that the probe will be of little use for either strain-specific or species-specific work. When a lower level of stringency washing was employed probe T37-5 hybridised to a larger selection of isolates including *Rt*162S7A, *Rt*JJD15, *Rv*1001, *Rv*1014, *Rm*2004 and TG1 (results not shown). This confirmed the earlier suspicions that the probe would be of no use for hybridisation studies of *Rhizobium*.

The overall conclusions that can be drawn from this study of strain-specific probes are that (i) it is possible to isolate strain-specific probes from RAPD profiles using the method of band-stab (Bjourson and Cooper, 1992) although a report by Bjourson *et al.*, (1992) of the use of subtraction hybridisation and DNA amplification suggests that this may be an easier method to rapidly identify strain-specific probes from smaller numbers of bacterial strains. (ii) In order to clearly identify strain-specific probes from the general background it is necessary to employ a high level of wash stringency. The need for this high level of stringency suggests that, at lower stringency levels, there is a certain degree of mismatching occurring between the bacterial genomes and the probes under investigation. This mis-matching indicates small DNA variations between the strains may have given rise to the strain-specific bands used here as strain-specific DNA probes. It also indicates that there are detectable variations within the *Rhizobium* genome which can be utilised for strain identification. (iii) Only one probe, T37-3, displays a high level of strain-specificity when hybridised to a selection of isolates. This probe, therefore, appears to be the only strain-specific probe identified from the *Rhizobium* and *Bradyrhizobium* isolates used in this study. (iv) However, the probe T37-2 produced results which suggest that it may be of use in studies involving a limited number of isolates. (v) The low number of true, strain-specific probes observed may be as expected within a bacterial genome where there is little 'excess' DNA, so that although small variations do occur, much of the DNA, within the whole population, will be conserved. (vi) These small variations in the genomic sequence can lead to the formation or loss of primer binding sites which may in turn lead to the production of apparent strain-specific bands in some isolates. (vii) Several of the probes, although not strain-specific, may be of use in experiments where the number of *Rhizobium* isolates under investigation is limited. (viii) Other non strain-specific probes (T1-2, T1-3, V1-2, T37-1) appeared of potential use as species-specific probes, although they were not tested further in this study. (ix) The result of some probes hybridising to the non-*Rhizobium* isolates indicates that a degree of homology exists between these isolates and the *Rhizobium* strains. This may be expected in some

instances as the non-*Rhizobium* isolates are also soil dwelling bacteria and would, therefore, presumably carry some similar genes to the *Rhizobium* strains. (x) The results described here, and those reported by Bjourson *et al.*, (1992), indicate that it is possible to identify DNA fragments, within *Rhizobium* isolates, which can be employed for strain-specific studies. This offers the potential of assessing *Rhizobium* strain interactions without the need to modify the genetic makeup of the isolates, i.e. by adding transposon or plasmid markers.

Tables and Figures

Table 6.1: This shows details of the strain-specific probes used in this study. Associated specific strains, the respective species and probe sizes are shown.

Probe	Species	Associated specific strain	Size (bp)
T1-1	*R. leg.* bv *trifolii*	JJD4	430
T1-2	*R. leg.* bv *trifolii*	RAC2	650
T1-3	*R. leg.* bv *trifolii*	RAC2	650
T1-4	*R. leg.* bv *trifolii*	RAC15	1000
V1-1	*R. leg.* bv *viciae*	1007	400
V1-2	*R. leg.* bv *viciae*	1023	775
M1-1	*R. meliloti*	2004	830
T37-1	*R. leg.* bv *trifolii*	RAC15	450
T37-2	*R. leg.* bv *trifolii*	JJD15	795
T37-3	*R. leg.* bv *trifolii*	JJD15	330
T37-5	*R. leg.* bv *trifolii*	RAC39	300

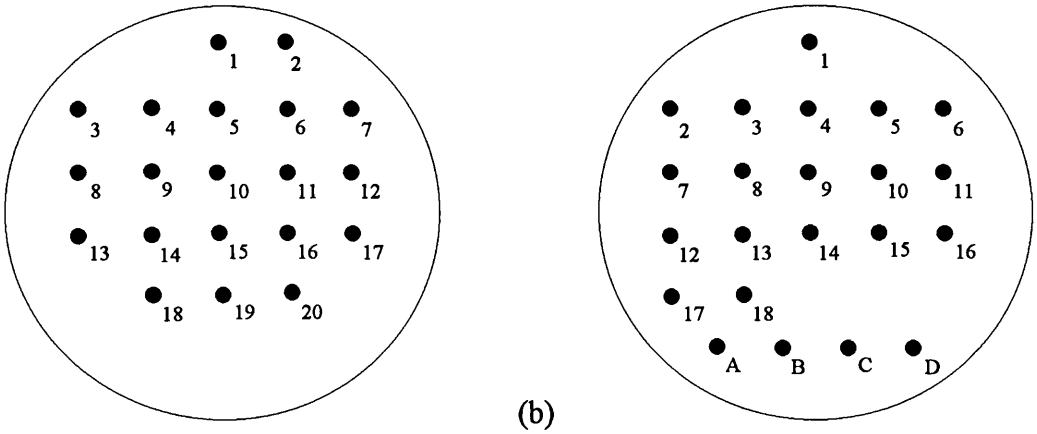


Figure 6.1: Isolate arrangements for probing with strain-specific probes (a) and species-specific probes (b). The isolates for strain-specific hybridisations were arranged numerically, 1-20, as follows: JJD4, RAC2, RAC15, RAC39, JJD15, (bv *trifolii*); 1007, 1014, 1023, (bv *viciae*); 3608, 3619, (bv *phaseoli*); 2004, 2009, (*R. meliloti*); 1001, (bv *viciae*); 3603, 3604, (bv *phaseoli*); 2000, 2001, (*R. meliloti*); TG1, TG3 & TG10 (Transgoed soil isolates). Isolates in (b) are arranged in numerical order as Table 2.1, (pg 46). Isolates A-D are controls, JJD4 (A) (bv *trifolii*); 1001 (B) (bv *viciae*); 2000 (C) (*R. meliloti*); Herring Sperm DNA (D).

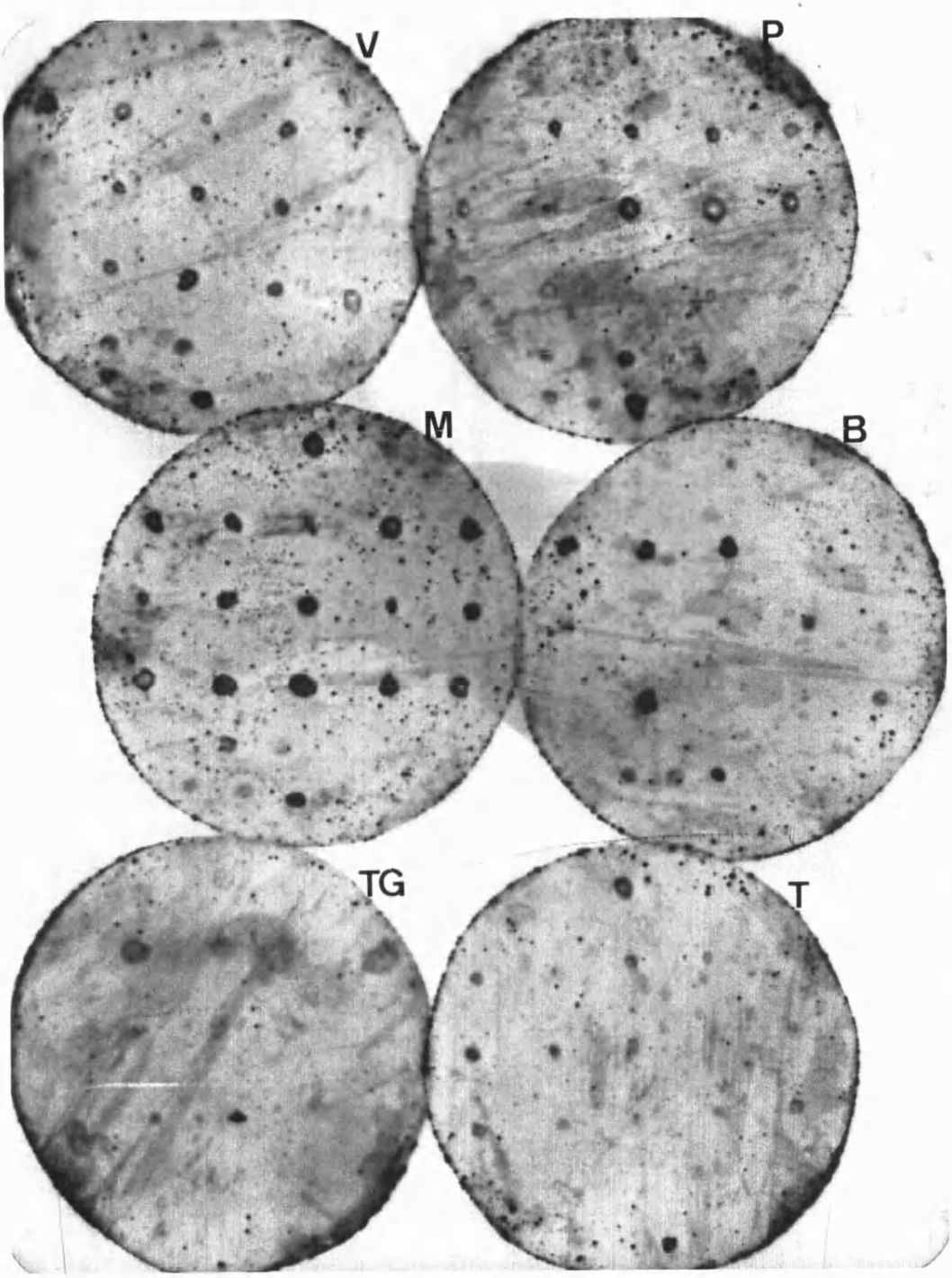


Figure 6.2: Results of hybridising colony blots with the species-specific probe M1A. Filters carry isolates from *R. leguminosarum* biovars *trifolii* (T), *viciae* (V), *phaseoli* (P), *R. meliloti* (M), *Bradyrhizobium* (B) and Transgoed soil isolates (TG).

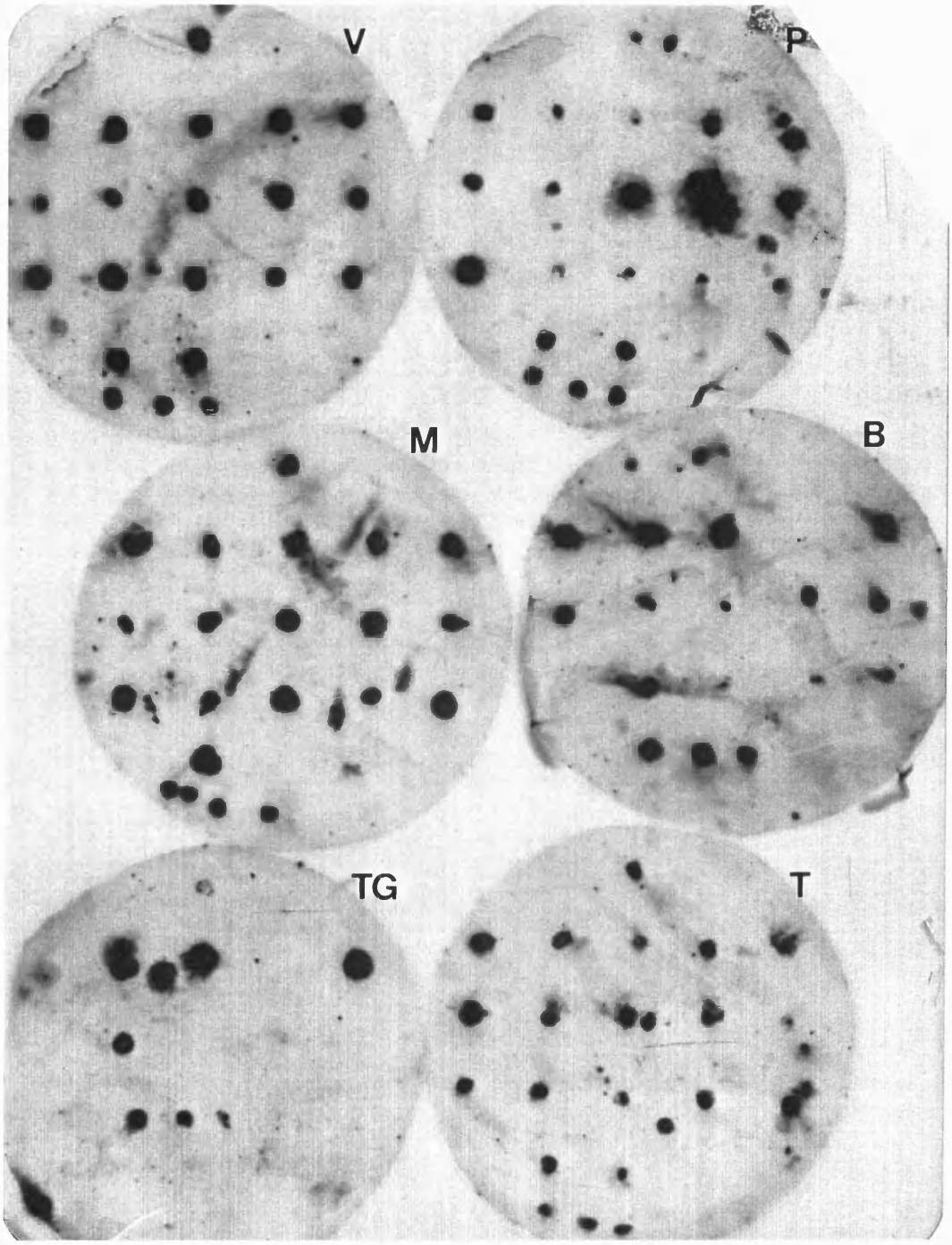


Figure 6.3: Results of hybridising colony blots with the species-specific probe M1B. Filters carry isolates from *R. leguminosarum* biovars *trifolii* (T), *viciae* (V), *phaseoli* (P), *R. meliloti* (M), *Bradyrhizobium* (B) and Transgoed soil isolates (TG).

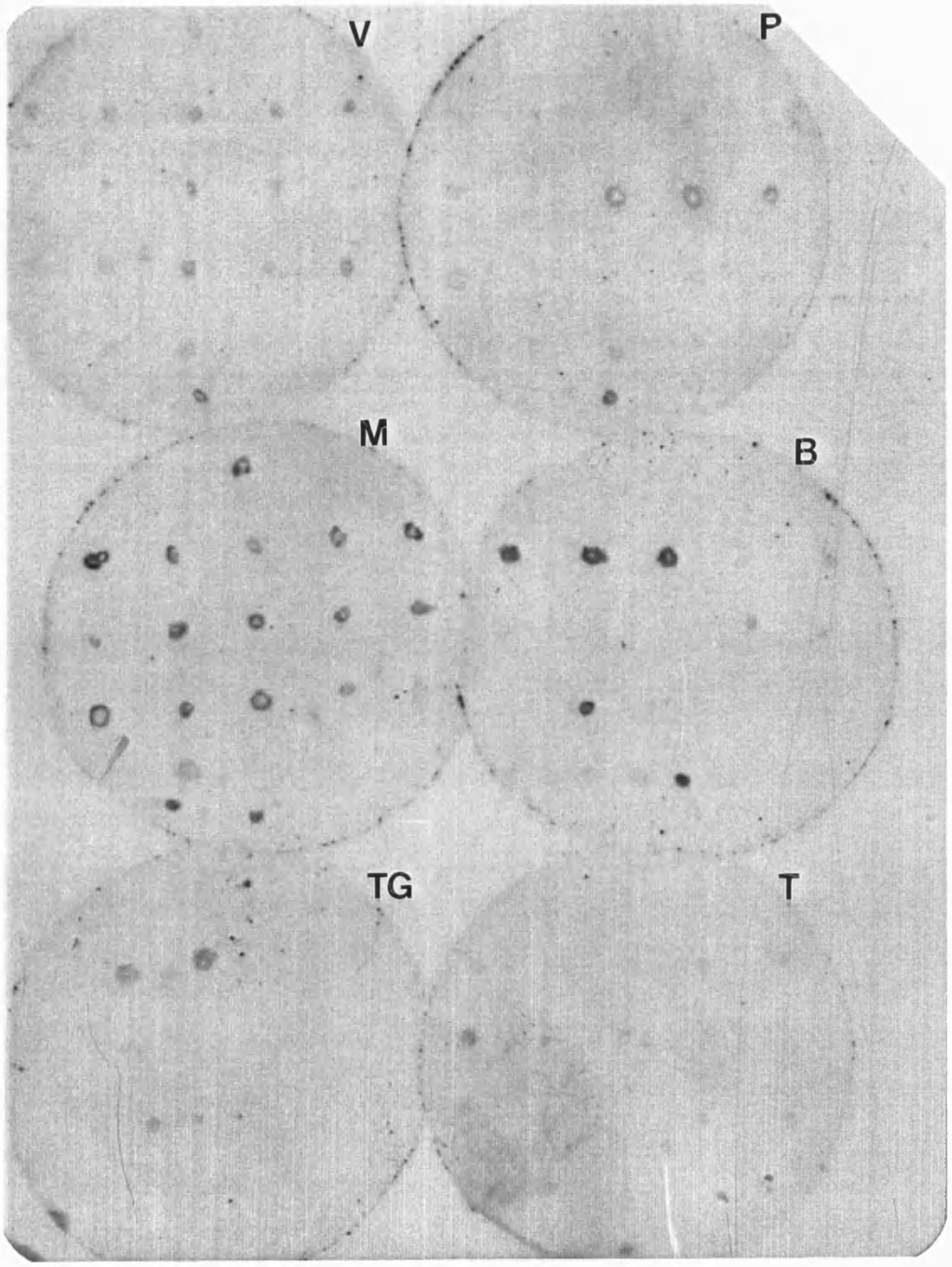


Figure 6.4: Results of hybridising colony blots with the species-specific probe M1C. Filters carry isolates from *R. leguminosarum* biovars *trifolii* (T), *viciae* (V), *phaseoli* (P), *R. meliloti* (M), *Bradyrhizobium* (B) and Transgoed soil isolates (TG).

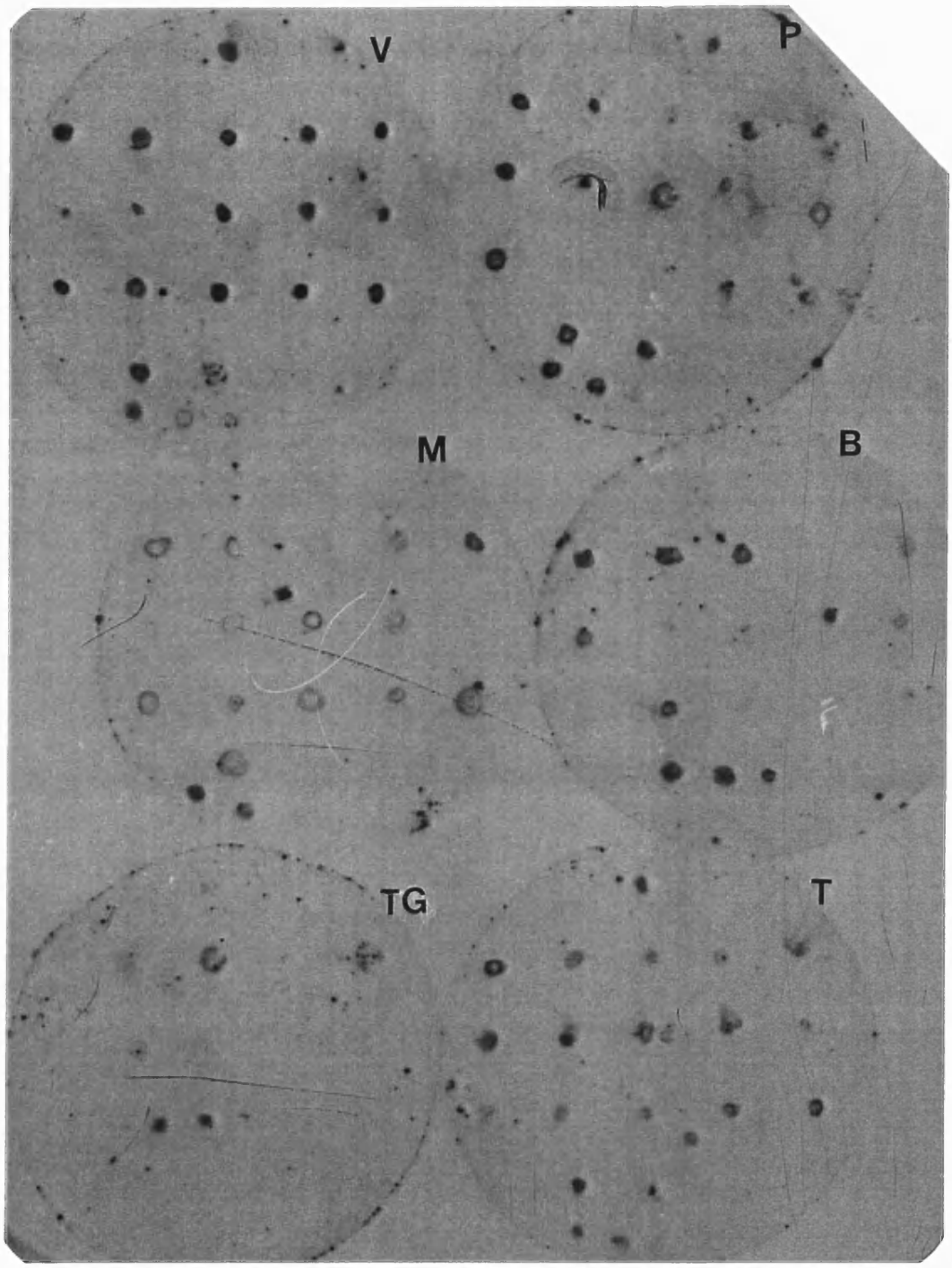


Figure 6.5: Results of hybridising colony blots with the species-specific probes V1D and T1E. Filters carry isolates from *R. leguminosarum* biovars *trifolii* (T), *viciae* (V), *phaseoli* (P), *R. meliloti* (M), *Bradyrhizobium* (B) and Transgoed soil isolates (TG).

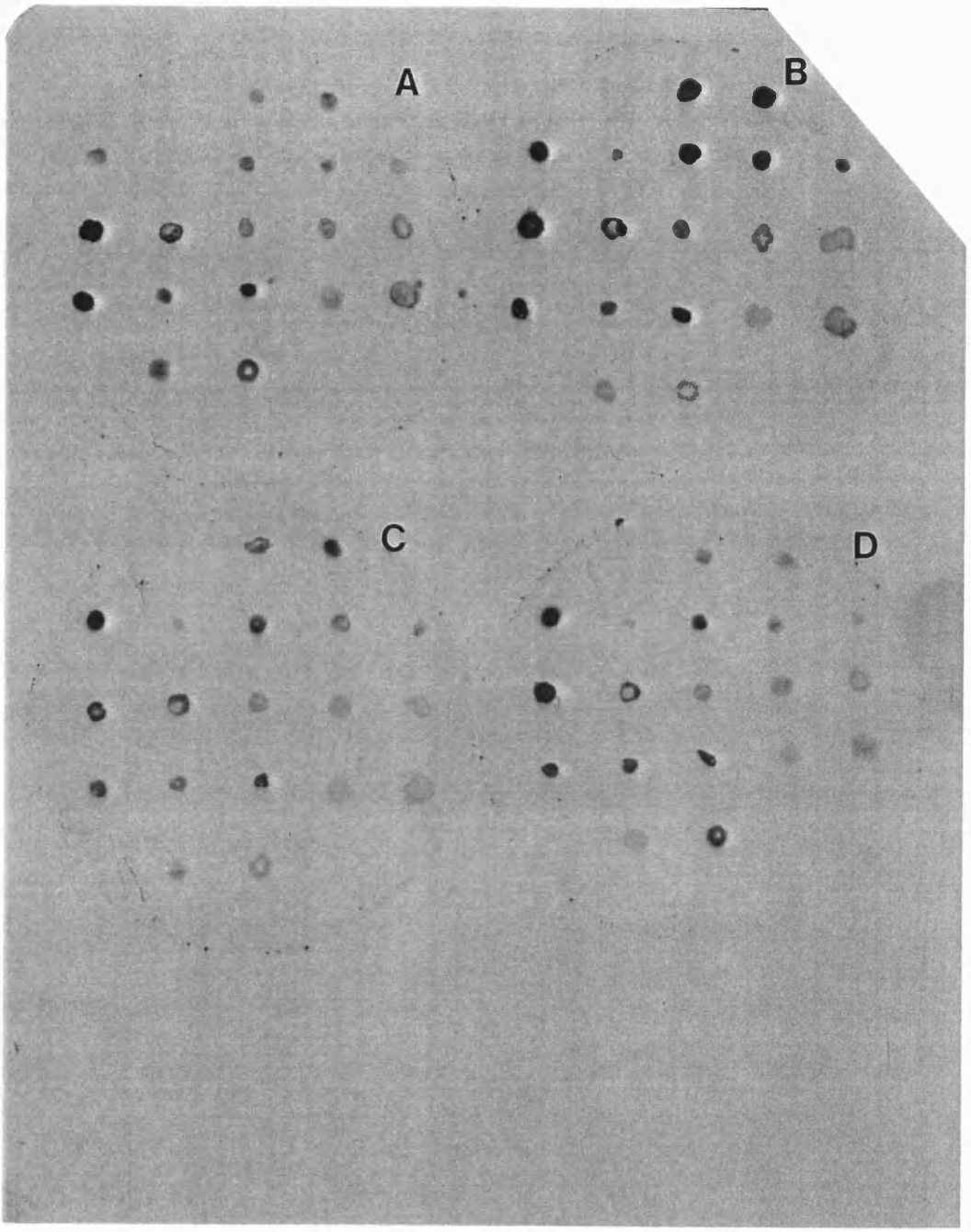


Figure 6.6: Results of hybridising colony blots with the strain-specific probes derived from amplification reactions using the primer SPH1. Filters were hybridised with probes T1-1 (A), T1-2 (B), T1-3 (C) and T1-4 (D).

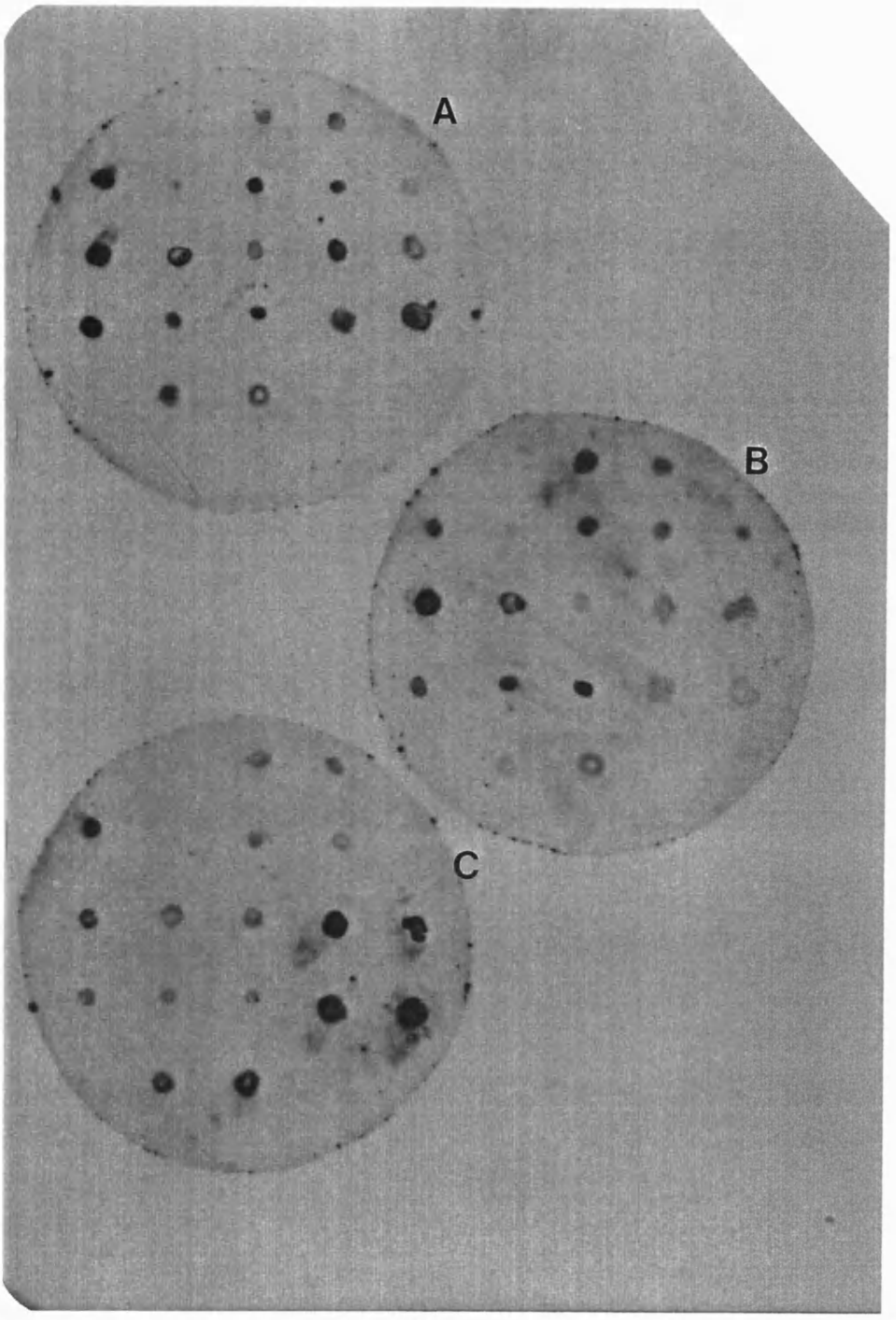


Figure 6.7: Results of hybridising colony blots with the strain-specific probes derived from amplification reactions using the primer SPH1. Filters were hybridised with probes V1-1 (A), V1-2 (B) and M1-1 (C).

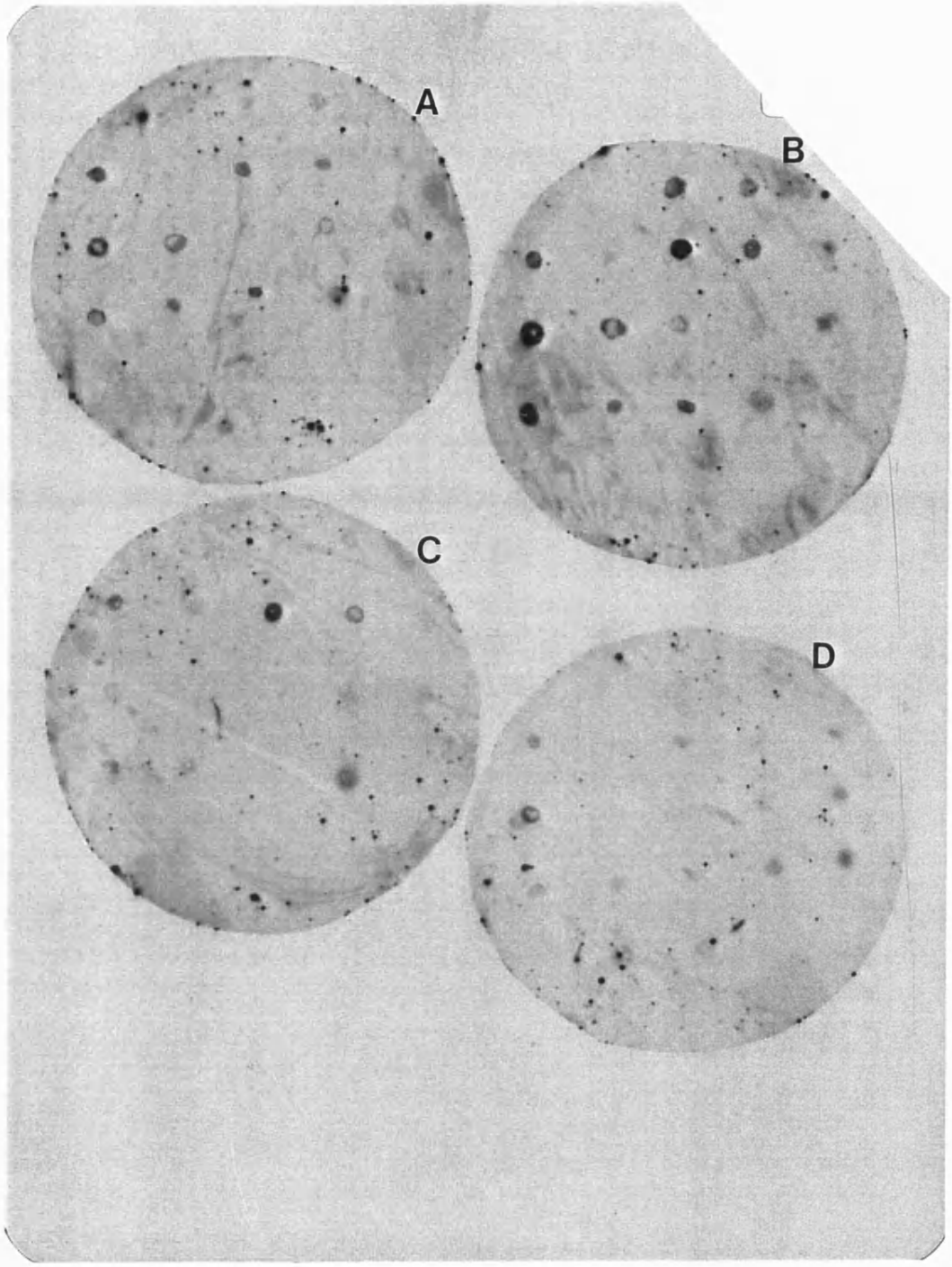


Figure 6.8: Results of hybridising colony blots with the strain-specific probes derived from dual primed amplification reactions using the primers SPH3 and SPH7. Filters were hybridised with probes T37-1 (A), T37-2 (B), T37-3 (C) and T37-5 (D).

An Assessment of the Distribution of RAPD-Derived Probes Within *Rhizobium* Isolates by Hybridisation to Restriction Digests of Total Genomic DNA.

Material from this chapter has been included in the paper:

DOOLEY J.J., HARRISON S.P., MYTTON L.R., DYE M., CRESSWELL A. & SKØT L. Characterisation of a Series of *Rhizobium* Species- and Strain-Specific DNA Probes. Submitted to FEMS Microbiology Letters.

Abstract

Hybridisation of a series of potential species-specific RAPD-derived probes to restriction digests of total genomic DNA from a selection of *Rhizobium* isolates revealed that the probes V1D and T1E, although derived from two separate biovars, reacted in an identical manner. This supports previous findings and indicates strongly that they are highly homologous. RFLP analysis also suggests that these two probes are species-specific. Results from hybridising the *R. meliloti* probes, M1A, M1B and M1C, also confirmed previous findings that these probes were not species-specific. However, RFLP profiles derived from probe M1B resulted in the classification of *R. meliloti* isolates into three sub-groups. Two of these groups, one and three, had similar RFLP patterns. Both groups comprised strains from the large *R. meliloti* cluster identified from RAPD profile analysis (Chapter 3; Dooley *et al.*, 1993). The third RFLP group comprised those isolates from the smaller RAPD cluster. It was also noted that the presence of a 4.2 Kb band, following hybridisation with probe M1B, was associated with isolates of the large RAPD cluster. A strong hybridisational signal from a 13.8 Kb band was also found to be indicative of membership of group one as defined by RFLP analysis.

Introduction

Bacterial classification can be achieved using a variety of techniques which have been discussed more fully in Chapter 1. These methods include those commonly known as fingerprinting of which RFLPs are an example.

The cleavage of total genomic DNA with restriction enzymes followed by gel electrophoresis, of the digested DNA, can be employed, on its own, for classification purposes. Mielenz *et al.*, (1979) used this method to differentiate between several *Rhizobium* species, whilst Laguerre *et al.*, (1992a) used this technique, in conjunction with plasmid profiling and RFLPs, to study *R. leguminosarum* bv *viciae*. The use of labelled DNA probes, however, provides a better method of resolving strain and species differences.

The type of probe employed varies and depends upon the particular interests of the research group. Insertion sequences (IS) have been used as probes for the identification of *R. meliloti* (Wheatcroft and Watson, 1987), whereas repeat sequences (RS) have been employed in studies of *Giardia intestinalis*, a parasitic protozoal species, (Ey *et al.*, 1992), *R. leguminosarum* bv *trifolii* (Schofield *et al.*, 1987) and *B. japonicum* (Minamisawa *et al.*, 1992). Restriction fragment length polymorphisms (RFLPs) have been used widely in the study of various bacterial species including *Lactobacillus helveticus* (de los Reyes-Gavilan *et al.*, 1992) and *Xanthomonas campestris* (Gabriel *et al.*, 1988) and the Mycoplasma like organism which causes X disease (Lee *et al.*, 1992). RFLPs have also been employed in the study of rhizobial species. Stanley *et al.*, (1985) used a selection of probes to create a series of RFLPs with which they identified two subgroups of *R. japonicum* (now *B. japonicum* [Elkan, 1992]). RFLPs have also been used, in conjunction with allozyme electrophoresis, to study *R. leguminosarum* bv *trifolii* (Demezas *et al.*, 1991). All three biovars of the species *R. leguminosarum* have been studied by Laguerre *et al*, (1992a,b) using RFLPs and plasmid profiling.

All these previous studies have employed specific probes to create RFLP patterns. In this study RAPD-derived probes, which have been described earlier (Chapters 3, 5 and 6), have been used to create RFLP patterns. This study was carried out in order to examine the distribution of these DNA fragments within each species and within individual strains.

Materials and Methods

All materials and methods used in this study are described in Chapter 2. Stringency washes, following hybridisation of species-specific probes were performed at 50° C in a wash solution of 1x SSC, 0.1% SDS. The strain-specific probe, T37-3, was subject to a stringency wash at 50° C in a wash solution of 0.1x SSC, 0.1% SDS.

Results and Discussion

***Probe M1A*:**

The results of hybridising DNA digests with the probe M1A are shown in Figures 7.1 to 7.3. At the level of stringency washing used the following observations were made.

R. leguminosarum bv *trifolii*:

When M1A was hybridised to digests of *R. leguminosarum* bv *trifolii* DNA no strains were seen to hybridise to the probe. This indicates little or no homology exists between the *R. meliloti* probe M1A and the species *R. leguminosarum* bv *trifolii*. A similar result was found when M1A was hybridised to RAPD profiles of this biovar. However, when colony blots were probed isolate *Rt*JJD4 was found to reveal a degree of homology to the probe although the signal intensity from this isolate was less then that from the control strain. It may, therefore, be that this colony blot result was not a true positive. The observations made from RFLP and RAPD analysis of strains of *R. leguminosarum* bv *trifolii* appear to suggest that M1A is of potential use as a species-specific probe.

R. leguminosarum bv *viciae*:

The results of hybridising M1A to restriction digests of DNA from *R. leguminosarum* bv *viciae* isolates are shown in Figure 7.1. The number of distinguishable bands seen on RFLPs from *R. leguminosarum* bv *viciae* isolates varies between zero (isolates *Rv*1004, *Rv*1012, *Rv*1014, *Rv*1015 and *Rv*1016) and ten (isolate *Rv*1025). The largest band observed is of size 11.4 Kb and the smallest of size 1.6 Kb. One band of size 2.5 Kb occurs in twelve of the isolates, *Rv*1001, *Rv*1011, *Rv*1013 and *Rv*1017-*Rv*1026, with varying intensities. The observed intensity differences could be due to a variation in copy number which suggests the presence of a plasmid borne

piece of DNA. Isolates *Rv*1017, *Rv*1025 and *Rv*1026 show the strongest signal with this probe while a very faint signal is seen in isolates *Rv*1021 and *Rv*1024. There appears to be no correlation between the presence of certain bands or band signal intensity and strain classification based on RAPD profiles. Isolates *Rv*1013, *Rv*1017 and *Rv*1025 reveal what may be a polymorphic band of size 2.8 Kb. This is about 300 bp larger than the 2.5 Kb band.

Isolate *Rv*1025 reveals three strongly hybridising bands of sizes 6.1 Kb, 3.2 Kb, and 2.5 Kb, the smaller of which is the common band just discussed. Very few of the remaining RFLP bands seen in the various isolates are shared by more than two isolates and many are found in one strain only.

These observations revealed that a certain degree of homology exists between isolates of *R. leguminosarum* bv *viciae* and the species *R. meliloti*. The results also suggest that the probe maybe homologous to a plasmid borne region of the genome as variations in intensity indicate variations in copy number. However, variations in copy number of such large plasmids would be rare and may therefore indicate that the differences could be caused by repeats of small segments of the genome. This tends to imply the presence of RSs of varying numbers within each strain. The absence of the band from some isolates again suggests it is not chromosomal, as this would make it universally present. However, these observations could also be the result of ISs which are present in some isolates but absent from others. Finally, the exchange of DNA between isolates of these two species has been reported (Djordjevic *et al.*, 1983; Broughton *et al.*, 1987). In these cases the material exchanged has been plasmid DNA which indicates that the bands observed on the DNA digest may be plasmid borne. The nature of transposable elements (transposons) means they move of from one strain to another and replicate themselves within each strain. This could also account for the variation in band intensity observed made here and also for the presence of the band in some strains while it is absent from others.

The differences in the hybridisation patterns between *R. leguminosarum* bv *trifolii* and *R. leguminosarum* bv *viciae*, observed with this probe, indicate that,

although closely related, variation between them does exist. It may therefore be possible to differentiate between some isolates from these two biovars using this probe.

When the stringency, of the washes, is increased the fainter bands, observed on the *R. leguminosarum* bv *viciae* filters, are removed leaving only the 3.9 Kb band observed in isolate *Rv*1007, two bands one of 2.5 Kb (the common band) and one of 2.3 Kb in isolate *Rv*1011, a band of 3.0 Kb in isolates *Rv*1017 and *Rv*1018 and the 3.2 Kb band in isolate *Rv*1025. When M1A was hybridised to RAPD profiles of strains of *R. leguminosarum* bv *viciae* and *R. leguminosarum* bv *trifolii* there were no bands apparent, which suggested that the probe maybe species-specific. However, cross-hybridisation with *R. leguminosarum* bv *phaseoli* isolates dispelled this theory and meant the probe was of little use for species-specific work. This observation was confirmed by colony blots where two isolates of *R. leguminosarum* bv *viciae* (*Rv*1004 and *Rv*1021) and one isolate of *R. leguminosarum* bv *trifolii* (*Rt*JJD4) displayed some cross-reactivity with the probe. It would be expected, however, that colony and RFLP results would be similar. The variations observed between RFLPs and colony blots may be the result of poor DNA fixation to the colony blot filters, or it may be that extra-cellular material, which remained following cell lysis, has shielded the DNA from the probe. Results obtained from both colony blots and RFLPs of *R. leguminosarum* bv *viciae* indicate that M1A is of little use as a species-specific probe.

R. leguminosarum bv *phaseoli*:

The results of hybridising M1A to digests of *R. leguminosarum* bv *phaseoli* DNA are shown in Figure 7.2. From Figure 7.2 it can be seen that very few bands are present in isolates of *R. leguminosarum* bv *phaseoli*. Only isolates *Rp*3603 and *Rp*3608 reveal the presence of two bands of sizes 11 Kb and 2.5 Kb. Isolates *Rp*3607, *Rp*3609, *Rp*3611 and *Rp*3614 display one band each of sizes 2.5 Kb, 2.5 Kb, 3.8 Kb and 2.2 Kb. respectively. These results indicate homology exists between the species *R. meliloti* and isolates of the species *R. leguminosarum* bv *phaseoli*. This, again, supports the hypothesis that M1A is of little use for species-specific work. The

presence of a 2.5 Kb band in some isolates is similar to the 2.5 Kb band seen in *R. leguminosarum* bv *viciae* isolates and may be from an identical region of the genome. This band was suggested as possibly being a plasmid borne fragment, which intimates plasmid exchange between strains of these two biovars, *phaseoli* and *viciae* has occurred. Isolates *Rv*1017, *Rp*3607, *Rp*3609 and *Rp*3614 all cluster together at the 80-100% level of similarity following analysis of RAPD profiles. The results of this Cluster Analysis and hybridisation with M1A indicate that isolate *Rv*1017 may have been mis-labelled as an isolate of *R. leguminosarum* bv *viciae*.

R. meliloti:

The results of self hybridisation between *R. meliloti* strains and the *R. meliloti* probe, M1A, are shown in Figure 7.3. The numbers of bands per strain varies between zero (5 isolates) and fourteen (isolate *Rm*2000). The largest band revealed by this probe is of size 20 Kb and the smallest 960 bp. Four bands of particular interest are those of sizes 8.9 Kb, 8.4 Kb, 7.7 Kb and 3.6 Kb. The 8.9 Kb band is seen in seven of the strains with a relatively high signal intensity in isolates *Rm*2001, *Rm*2003 and *Rm*2007. A polymorphic band of slightly smaller size (8.5 Kb) can be seen in isolates *Rm*2000, *Rm*2007 and *Rm*2009, the latter strain not displaying any presence of the 8.9 Kb band. The 8.4 Kb band is seen in only two isolates, *Rm*2006 and *Rm*2015. This band may, however, be a second polymorph of the 8.9 Kb band as both these strains also contain this larger band. The intensities of these two bands (8.9 Kb and 8.4 Kb) suggest a higher copy number of the smaller band. The third band (7.7 Kb) is also seen in relatively high amounts, based on signal intensity, in the strains *Rm*2001, *Rm*2003, *Rm*2007, *Rm*2009 and *Rm*2011, the latter isolate showing a very strong signal with the probe. No other isolates contain this band although there are possible larger polymorphic bands in isolates *Rm*2000, *Rm*2008, *Rm*2013 and *Rm*2015 and smaller polymorphic bands in isolates *Rm*2012 and *Rm*2015. These possible polymorphic bands are very weak in intensity when compared to the 7.7 Kb sized band. All the isolates, with the exception of *Rm*2012 and *Rm*2015, displaying the presence of the 7.7

Kb band and its potential polymorphic bands are found in the larger cluster formed when RAPD profiles are analysed (Chapter 3). This result therefore appears to confirm the finding that *R. meliloti* may be composed of two sub-groups as described in previous reports (Chapter 3; Young *et al.*, 1985; Eardly *et al.*, 1990; Dooley *et al.*, 1993). The smallest band of interest (3.6 Kb) can be seen in strains *Rm*2000, *Rm*2001, *Rm*2003, *Rm*2006, *Rm*2008, *Rm*2010, *Rm*2011 and *Rm*2015. A strong signal is seen in isolates *Rm*2001, *Rm*2003 and *Rm*2011 but weaker signals in the other isolates. Several other weak signals can be seen in various isolates and there is also a strong band of size 960 bp. in isolates *Rm*2000 and *Rm*2001 only.

The overall banding patterns observed following hybridisation of *R. meliloti* filters with M1A (Figure 7.3) are not of great value for strain classification. However, the presence of certain bands in particular strains does correspond to isolate clustering when RAPD profiles were analysed and is similar to the presence of certain bands when these RAPD profiles were hybridised with M1A.

Probe M1B:

R. leguminosarum bv *trifolii*:

When M1B was hybridised to DNA digests from isolates of *R. leguminosarum* bv *trifolii* no banding patterns were obtained. This result is similar to that observed when M1A was employed. When probes M1A and M1B were hybridised to RAPD profiles from *R. leguminosarum* bv *trifolii* isolates no hybridisation was observed, which is in agreement with results found here. However, when M1B was hybridised to colony blots nearly all isolates from all species, including *R. leguminosarum* bv *trifolii*, were seen to hybridise to the probe. This indicated that M1B had some homology with nearly every soil-borne bacterial strain examined. The results presented here, however, intimate that there is little homology between the *R. leguminosarum* bv *trifolii* isolates and M1B.

R. leguminosarum bv *viciae*:

The results of hybridising *R. leguminosarum* bv *viciae* filters with M1B are shown in Figure 7.4. The number of visible bands per strain varies between zero and five, with five isolates (*Rv*1004, *Rv*1012, *Rv*1014, *Rv*1015 and *Rv*1016) displaying no pattern and two (*Rv*1023 and *Rv*1025) revealing five bands. Five intense bands can be seen in isolates *Rv*1007, *Rv*1011, *Rv*1017, *Rv*1018 and *Rv*1025. These bands are of sizes 3.9 Kb (isolate *Rv*1007), 2.5 Kb (strain *Rv*1011), 3.0 Kb (isolates *Rv*1017 and *Rv*1018) and 4.3 Kb and 3.2 Kb (isolate *Rv*1025). Bands of these sizes are also observed when M1A is hybridised to the digested genomic DNA of *R. leguminosarum* bv *viciae* isolates, although the relative intensities of the signals varies. When RAPD profiles were probed with M1B no bands were observed although colony blots did reveal a high level of cross-hybridisation. The lack of hybridisation to RAPD profiles is probably owing to either a variation in the primer binding sites which has resulted in no amplification of this part of the genome, from *R. leguminosarum* bv *viciae* isolates, or there has been a change in the distance between the primer sites (owing to an insertion) which has resulted in the maximum distance for DNA amplification (about 3.0 Kb) being exceeded.

R. leguminosarum bv *phaseoli*:

Following hybridisation of M1B to DNA digests of *R. leguminosarum* bv *phaseoli* isolates the results shown in Figure 7.5 were produced. Band sizes seen on the *R. leguminosarum* bv *phaseoli* filter vary between 25.7 Kb and 2.3 Kb, with only ten isolates displaying any hybridisation with the probe. Isolates *Rp*3603, *Rp*3607, *Rp*3608 and *Rp*3609 show nearly identical patterns with four common bands of sizes 11.2 Kb, 2.8 Kb, 2.5 Kb and 2.2 Kb. This suggests that these four isolates are from the same sub-group which are presently believed to comprise *R. leguminosarum* bv *phaseoli* (Pinero *et al.*, 1988; Martinez-Romero *et al.*, 1991; Dooley *et al.*, 1993). When the results of RAPD analysis are examined (Chapter 3) the strains *Rp*3607, *Rp*3608 and *Rp*3609 are found to be grouped together in a cluster formed at the 80-

100% similarity level. Strain *Rp*3603, however, is found to associate with the *R. leguminosarum* bv *trifolii* / *viciae* cluster formed at the 75-80% level of similarity. This observation appears to contradict the findings made from RFLP analysis and may indicate the existence of variations between these strains which are detectable following DNA amplification but not after DNA hybridisation studies. This would, therefore, suggest that the differences are in the primer binding regions which alters the RAPD pattern observed in isolate *Rp*3603. The 2.8 Kb band displays an intense signal in these isolates and is also present, as the sole band in isolate *Rp*3617. The 2.5 Kb and 2.2 Kb bands are found as a couplet in strains *Rp*3603, *Rp*3607 and *Rp*3608. The smaller band is found alone in isolates *Rp*3609, *Rp*3611, *Rp*3614 and *Rp*3626. Bands of size 2.5 Kb and 2.2 Kb are also found following hybridisation of these digests with M1A. The presence of these bands following hybridisation with both probes intimates at a degree of homology existing between probes M1A and M1B. The remaining bands of various sizes are found in only one or two strains. Also, all these bands, with the exception of a 4.6 Kb band found only in isolate *Rp*3611, display a poor signal intensity which is eliminated with a stronger stringency wash.

R. meliloti:

The results of probing *R. meliloti* isolates with M1B are shown in Figure 7.6. Five of the *R. meliloti* strains do not show any banding pattern with M1B while those isolates with bands have between three and fourteen distinguishable bands. The largest of these bands is 20 Kb and the smallest 3.4 Kb in size. This small band appears to be polymorphic and is discussed further below.

Three sets of isolates with identical banding patterns can be seen on the filters; group one consists of isolates *Rm*2001, *Rm*2003, *Rm*2007 and *Rm*2009, group two of strains *Rm*2006 and *Rm*2015 and group three of isolates *Rm*2000, *Rm*2008, *Rm*2010 and *Rm*2013 (although the strain *Rm*2013 lacks some bands from the pattern). This third cluster, however, would link with the first set if some of the intermediate bands were removed. This suggests a sub-division may exist within one of the groups so that

the *R. meliloti* species comprises two major strain clusters of which one of these groups comprises two smaller sets. When clusters formed by RAPD and RFLP analysis are examined the strains comprising group two (from the RFLP analysis) are found to comprise the smaller RAPD cluster formed at the 80-100% level of similarity (Chapter 3). The two remaining strains forming this RAPD group are *Rm*2002 and *Rm*2012, however, no RFLP patterns have been produced for these strains. The larger RAPD cluster (Chapter 3) comprises those strains forming groups one and three from RFLP analysis. This observation confirms the findings from Chapter 3 where a subdivision of the species *R. meliloti* was found following analysis of RAPD patterns. This sub-division of the species *R. meliloti* has also been suggested previously (Young, 1985; Eardly *et al.*, 1990; Dooley *et al.*, 1993). No RFLP pattern was produced for the strain *Rm*2017 which from RAPD analysis was found to be mis-labelled (Chapter 3). RAPD analysis suggested strain *Rm*2017 was an isolate of *R. leguminosarum* bv *phaseoli*, but without an RFLP pattern this cannot be confirmed here. These observations indicate that M1B maybe of use for differentiating between isolates of the two sub-groups of the species *R. meliloti*.

From an observation of the RFLP patterns five bands of sizes 13.8 Kb, 5.3 Kb, 4.2 Kb, 3.5 Kb and 3.3 Kb, which are fairly common to the strains of *R. meliloti*, can be seen. The 13.8 Kb band is seen in nine isolates *Rm*2000, *Rm*2001, *Rm*2003, *Rm*2004, *Rm*2007, *Rm*2009, *Rm*2010, *Rm*2011 and *Rm*2013 with varying signal intensities. The strongest signal is associated with those isolates comprising group one as described above, i.e. strains *Rm*2001, *Rm*2003, *Rm*2007 and *Rm*2009. Relatively weak signals are seen in the other isolates which are mostly group three strains with the addition of isolates *Rm*2004 and *Rm*2011. All the isolates are from the larger cluster formed at the 80-100% level of similarity following the application of Cluster Analysis to the RAPD profiles. This may indicate that presence of this 13.8 Kb band is indicative of membership of this subgroup of the *R. meliloti* species and so may be of use as a marker for such purposes. The 5.3 Kb band is seen in the same isolates as the 13.8 Kb band with the exception of isolates *Rm*2004 and *Rm*2013 which lack the

smaller band. The strains *Rm*2006 and *Rm*2015 contain the 5.3 Kb band but not the larger 13.8 Kb band. These two strains are from the smaller *R. meliloti* subgroup formed after analysis of RAPD profiles. The presence of the 5.3 Kb band in both subgroups of *R. meliloti* therefore indicates a degree of homology exists between both these subgroups. The third smallest band (4.2 Kb) is present in eight isolates which form groups one and three as described above. This again confirms the earlier findings about a subdivision within the species *R. meliloti* and indicates that the presence of this band maybe indicative of membership of the sets one or three as defined by RFLP analysis. The remaining two bands, of sizes 3.5 Kb and 3.3 Kb, are the most common appearing together or alone in all twelve isolates which show banding patterns with this probe. The occurrence of these bands suggests that they are polymorphic. They appear as a doublet in strains *Rm*2000, *Rm*2001, *Rm*2003, *Rm*2010 and *Rm*2011. The larger band (3.5 Kb) is present alone in isolates *Rm*2004, *Rm*2006, *Rm*2008, *Rm*2013 and *Rm*2015 and the smaller band (3.3 Kb) is present in strains *Rm*2007 and *Rm*2009. The smaller of these bands is therefore only present in isolates from the larger subgroup identified following RAPD profile analysis.

Overall observations made with M1B suggest that it is of little use as a species-specific marker as it displays high levels of cross-reactivity with all species except *R. leguminosarum* bv *trifolii*. However, when used to back probe to restriction digests of *R. meliloti* genomic DNA this probe appears to be of use in classifying these strains into the two major sub-groups identified from RAPD profile analysis. This probe also allows for a further sub-division of the larger cluster, identified by RAPD profile analysis, into two smaller sets. It would be interesting to use this probe to analyse restriction digests of *Bradyrhizobium* digests in order to ascertain if this probe can be used to differentiate the two clusters formed following RAPD profile analysis.

Probe M1C:

R. leguminosarum:

No bands were present in strains of *R. leguminosarum* bvs *trifolii* and *phaseoli* when probed with M1C suggesting that this probe may be species-specific. However, four isolates of *R. leguminosarum* bv *viciae* reveal RFLP banding patterns with M1C (Figure 7.7) so dispelling this theory. Two of these isolates, *Rv*1017 and *Rv*1019, have only a single band each, of sizes 2.7 Kb and 3.0 Kb respectively. The remaining two isolates, *Rv*1014 and *Rv*1015, have five and six bands respectively. Isolate *Rv*1014 has bands of size 4.7 Kb, 4.6 Kb, 3.5 Kb, 2.0 Kb and 1.8 Kb and strain *Rv*1015 has bands of size 5.5 Kb, 5.1 Kb, 4.6 Kb, 4.4 Kb, 3.6 Kb and 3.0 Kb. Only the small (1.8 Kb) band in isolate *Rv*1014 displays a strong signal with this probe at this level of stringency washing. When greater levels of stringency washing are used this small band is the only signal seen from the *R. leguminosarum* bv *viciae* isolates. This, however, still indicates the existence of a degree of homology between these two species. These observations vary from those made when RAPD profiles were hybridised with M1C. Southern blots of RAPD profiles revealed only the presence of a single 630 bp. band in *R. leguminosarum* bv *phaseoli* isolate *Rp*3618. The lack of a band in this isolate following hybridisation to DNA digests is probably owing to the increased stringency wash conditions which will remove any unstable probes. The presence of bands, following RFLP analysis, in isolates of *R. leguminosarum* bv *viciae* which failed to display there presence following RAPD analysis indicates that this part of the genome has not been amplified during the RAPD reaction.

R. meliloti:

The results of hybridising M1C to DNA digests of *R. meliloti* are shown in Figure 7.8. From this figure it can be seen that band numbers vary between zero (isolates *Rm*2002, *Rm*2004, *Rm*2005, *Rm*2008, *Rm*2012, *Rm*2016 and *Rm*2017) and ten (isolate *Rm*2002). Band sizes vary between 15.1 Kb (isolate *Rm*2007) and 1.1 Kb (isolate *Rm*2001).

From a general observation of the results a group of isolates with similar banding patterns is seen to emerge. This cluster comprises strains *Rm*2001, *Rm*2003 and *Rm*2007 and possibly isolates *Rm*2009 and *Rm*2011 which have slight variations in their profiles. From RAPD profile analysis (Chapter 3) all these isolates are found in the larger *R. meliloti* cluster formed at the 80-100% level of similarity. These strains are also part of sub-group one, as defined by RFLP analysis using M1B. The six major bands which are shared amongst these isolates are of sizes 11.9 Kb, 8.5 Kb, 4.7 Kb, 4.4 Kb, 3.5 Kb and 2.8 Kb. The variations that can be seen include the replacement of the 8.5 Kb band with a smaller 7.6 Kb band in isolates *Rm*2003 and *Rm*2011 and a 3.8 Kb band in place of the 3.5 Kb band in isolate *Rm*2009. Isolates *Rm*2009 and *Rm*2011 also lack the 4.4 Kb band while the latter strain has a 2.9 Kb band in place of the 2.8 Kb band. The variations between patterns indicate that although highly related (>80% similarity) these isolates do have genomic differences which can be detected using RFLPs.

The observations made with both probes M1C and M1B indicate that there maybe several sub-groups within the species *R. meliloti*. This is similar to findings reported previously (Young, 1985; Eardly *et al.*, 1990; Dooley *et al.*, 1993). The results presented here also indicate that M1C maybe of use for identification of strains composing the large cluster as defined by RAPD analysis (Chapter 3). This probe is, however, inefficient at producing useful profiles from the other isolates and means they cannot be linked based on their RFLP patterns.

Probes V1D and T1E:

Both these probes reacted in an identical manner and have therefore been dealt with simultaneously. This observation is similar to that made when these probes were hybridised to blots of RAPD profiles and colony blots. This, therefore, again confirms the high level of homology between these two probes.

R. leguminosarum bv *trifolii*:

The results of hybridising probes V1D or T1E to restriction digests of DNA from isolates of *R. leguminosarum* bv *trifolii* are shown in Figure 7.9. From observations of Figure 7.9 it can be seen that all the isolates, with the exception of strain *Rt*221 (lane 5), reveal the presence of at least one band with this probe. The largest distinguishable band, of size 3.7 Kb, is found in isolates *Rt*0404 and *Rt*162P17. The smallest band, of size 1.5 Kb, is seen in four isolates, *Rt*5, *Rt*7D5, *Rt*R1.3 and *Rt*RAC41. A common band of size 2.26 Kb can be seen in all isolates (except strain *Rt*221 which reveals no banding pattern) including isolates *Rt*35 and *Rt*JJD17, which display very weak signals at this band size. The presence of, in effect, a single band in the RFLPs profiles of the *R. leguminosarum* bv *trifolii* isolates is similar to results obtained when RAPD blots were back-probed with this probe (Chapter 5). This suggests the piece of DNA amplified during the RAPD analysis and used to construct either of the probes is a single copy DNA fragment. The presence of this fragment in all isolates of *R. leguminosarum* bv *trifolii*, whether hybridised with V1D or T1E, indicates a high degree of homology exists between strains of *R. leguminosarum* bv *trifolii* and both the probes. The results also indicate that the DNA is from a highly conserved region of the genome otherwise a greater amount of variation in band numbers and size would be expected.

Variations between the band patterns from the isolates do exist. These include a 2.3 Kb band, which is possibly polymorphic to the 2.26 Kb common band, which is found in isolates *Rt*0404, *Rt*162BB1, *Rt*162P17 and *Rt*162X7a. A slightly larger band

of about 2.33 Kb, which again maybe polymorphic, can be seen in isolates *Rt*3, *Rt*35 and *Rt*162S7a. Isolates *Rt*162S7a and *Rt*162BB1 also display the presence of a band of size 2.5 Kb. These small, possibly polymorphic, band variations indicate that differences between the *R. leguminosarum* bv *trifolii* isolates do exist.

Several isolates also appear to have identical band patterns which indicates that sub-sets of isolates exist within this biovar. The three groups that can be identified include isolates *Rt*JJD4 and *Rt*JJD9, isolates *Rt*5, *Rt*7D5, *Rt*R1.3 and *Rt*RAC41 and isolates and *Rt*0404, *Rt*162P17 and *Rt*162X7a. Strains *Rt*3, *Rt*162S7a and *Rt*162BB1 all have unique profiles not seen in any other isolate. From results of RAPD analysis it was found that strains *Rt*0404, *Rt*162P17 and *Rt*162X7a associate at the 75-80% level of similarity. They also all fall spatially within the same region of the cluster. Isolates *Rt*JJD4 and *Rt*JJD9 also lie close to each other within the same cluster formed at the 75-80% level of similarity. The other isolates, however, are divided between the two clusters.

R. leguminosarum bv *viciae*:

The results of probing DNA digests from *R. leguminosarum* bv *viciae* isolates with V1D or T1E are shown in Figure 7.10. From Figure 7.10 it can be seen that band numbers, per strain, vary between one (isolate *Rv*1007) and thirteen isolate (*Rv*1026). The band sizes vary between a minimum of 1.8 Kb and a maximum of 9.9 Kb, both of which are in isolates *Rv*1025 and *Rv*1026. A band of 3.6 Kb is apparent in at least nine of the isolates (strains *Rv*1017 to *Rv*1026) and, allowing for gel distortions, is probably also present in isolates *Rv*1001 and *Rv*1013. This 2.6 Kb band maybe equivalent to the 2.3 Kb band seen in *R. leguminosarum* bv *trifolii* isolates, although it is about 300 bp larger. The variation of 300 bp between these two bands may, therefore, be of use for separating isolates from the two biovars. The 2.6 Kb band is not, however, universally distributed throughout the *R. leguminosarum* bv *viciae* isolates, it being replaced by two polymorphic bands of 2.8 Kb and 2.5 Kb in strain *Rv*1011 and a band of 3.6 Kb in isolate *Rv*1007. Several isolates, especially *Rv*1001, *Rv*1025 and *Rv*1026, also display

the presence of a selection of larger bands of varying sizes. This variation in the profiles suggests that there is a greater amount of genomic variation amongst strains of *R. leguminosarum* bv *viciae* than there is amongst *R. leguminosarum* bv *trifolii*. However, none of the *R. leguminosarum* bv *viciae* profiles are of use for classifying the isolates as no two strains appear to have the same RFLP profile. This tends to suggest that there are few highly related isolates in the species *R. leguminosarum* bv *viciae*. This is similar to results from RAPD profile analysis (Chapter 3) where *the R. leguminosarum* bv *trifolii* and *R. leguminosarum* bv *viciae* isolates formed clusters at the 75-80% level of similarity. The other species had strains which clustered at the 80-100% level of similarity. These observations vary from those found when V1D or T1E were hybridised to the *R. leguminosarum* bv *trifolii* DNA digests. The *R. leguminosarum* bv *trifolii* strains all have very similar RFLP profiles and the small variations between them are useful for a limited degree of strain classification.

R. leguminosarum bv *phaseoli*:

Probing restriction digests of *R. leguminosarum* bv *phaseoli* with V1D or T1E produced the results shown in Figure 7.11. Observations of Figure 7.11 reveal that eleven strains display banding patterns with this probe. The bands vary in size between 23.4 Kb and 1.6 Kb and the number per strain varies between one (isolates *Rp*3613 and *Rp*3615) and nine (isolates *Rp*3603 and *Rp*3608). Four bands of sizes 10.7 Kb, 2.6 Kb, 2.4 Kb and 2.3 Kb are seen in up to eight of the strains of *R. leguminosarum* bv *phaseoli*. The smallest of these bands is the same size as the common band seen in the *R. leguminosarum* bv *trifolii* isolates. This may suggest that this band is shared by both *R. leguminosarum* bv *phaseoli* and *R. leguminosarum* bv *trifolii* isolates. However, the presence of this DNA is not consistent within the *R. leguminosarum* bv *phaseoli* isolates with only eight isolates (*Rp*3603, *Rp*3606, *Rp*3607, *Rp*3608, *Rp*3609, *Rp*3611, *Rp*3617 and *Rp*3626) from this study displaying the presence of this DNA fragment. The presence of the band in these isolates may, therefore, indicate that they are more related to the *R. leguminosarum* bv *trifolii* isolates than those *R. leguminosarum* bv

phaseoli isolates which do not contain the band. From RAPD results (Chapter 3) isolate *Rp*3603 was found to associate with the *R. leguminosarum* bv *trifolii / viciae* cluster formed at the 75-80% level of similarity. The remaining strains were mostly linked into the large *R. leguminosarum* bv *phaseoli* group formed at the 80-100% level of similarity. This indicates a high level of homology exists between these isolates. The presence of the band in some isolates and not in others highlights a division within the *R. leguminosarum* bv *phaseoli* group. Similar reports have been made previously (Chapter 3; Pinero *et al.*, 1988; Martinez-Romero *et al.*, 1991; Dooley *et al.*, 1993).

Examination of the banding patterns from isolates in Figure 7.11 suggests that the strains *Rp*3603, *Rp*3607, *Rp*3608 and *Rp*3617 (although this latter strain has faint band signals) form a group distinct from the other strains which display banding patterns. When the results from analysis of the RAPD fingerprints are examined the strains *Rp*3607 and *Rp*3608 are found to cluster within the ten strain cluster seen at the 80-100% level of similarity. The remaining strains from this RAPD cluster are *R. leguminosarum* bv *phaseoli* isolates *Rp*3606, *Rp*3609, *Rp*3614, *Rp*3624, *R. leguminosarum* bv *trifolii* isolate *Rt*RAC35, *R. leguminosarum* bv *viciae* isolates *Rv*1015 and *Rv*1017 and *R. meliloti* isolate *Rm*2017. The *R. leguminosarum* bv *phaseoli* isolates *Rp*3606, *Rp*3609 and *Rp*3614 all reveal the presence of the 2.3 Kb band. The presence of this band in these strains and their clustering with isolates of *R. leguminosarum* bvs *trifolii* and *viciae* when RAPD profiles are analysed suggests a certain degree of homology exists between these particular strains and the species *R. leguminosarum*. The isolate *Rp*3603 fails to cluster with any *R. leguminosarum* bv *phaseoli* when RAPD profiles are analysed, however, it does cluster with the upper *R. leguminosarum* bv *trifolii / viciae* cluster formed at the 75-80% similarity level. The results from RFLP analysis, therefore, appear to lend a degree of support to the findings obtained from RAPD profile analysis. It may be that the strains showing RFLP patterns with these probes are either *R. leguminosarum* bv *phaseoli* Group I isolates or are of the recently formed species *R. tropici* as defined by (Martinez-Romero *et al.*, 1991).

The remaining isolates displaying RFLP patterns are *Rp*3613 and *Rp*3615 which carry only a single band of size 3.1 Kb. These two strains also have identical RAPD profiles when amplified with SPH1. Isolate *Rp*3626 is the final strain to reveal any homology with V1D or T1E. This strain clusters with isolates *Rp*3618 and *Rp*3619 when RAPD patterns are analysed, however, no banding patterns are produced, for these two strains, with RFLP analysis so no comparison can be made.

The partial patterning of the strains from *R. leguminosarum* bv *phaseoli* supports previous findings (Chapter 3; Pinero *et al.*, 1988; Dooley *et al.*, 1993) that the species *R. leguminosarum* bv *phaseoli* is not a homologous group but is a compilation of assorted strains which have the ability to nodulate beans (*Phaseolus vulgaris*).

R. meliloti:

No pattern was obtained following hybridisation of V1D or T1E to the DNA digests of *R. meliloti* isolates. This indicates a lack of any homology between these *R. leguminosarum* probes and the species *R. meliloti* and indicates that the probes may be of use for species-specific studies. This observation is in accordance with that made when RAPD profiles were hybridised to this probe. However, when colony blots were probed at least one isolate (*Rm*2005) indicated a degree of homology with V1D and T1E. The intensity of the signal from this isolate was, however, not quite as strong as that from the control isolate so that a false positive may have been made.

RFLP profiles obtained following hybridisation of the species-specific probes M1A, M1B, M1C, V1D and T1E to DNA digests confirm, in the most part, findings made in previous chapters. The following general points of interest can, however, be drawn from the hybridisation results. (i) The three *R. meliloti* derived probes are of little use as species-specific probes as they all display varying levels of cross-hybridisation to *R. leguminosarum* isolates. (ii) However, the profiles obtained following hybridisation of M1B to DNA digests of *R. meliloti* isolates appears to be more useful for classification of this species than previous studies (Chapters 5 and 6)

have indicated. These profiles intimate the existence of three sub-sets within the species *R. meliloti*, although two of these groups are closely related. This result is similar to those seen following RAPD profile analysis (Chapter 3) where two clusters were in evidence at the 80-100% level of similarity. The three sub-groups identified from RFLP analysis were found to be in accordance with the RAPD clusters. These results are similar to those reported previously (Young, 1985; Eardly *et al.*, 1990; Dooley *et al.*, 1993). (iii) It was also found that certain RFLP bands observed on filters hybridised with M1B coincide with membership of the sub-groups described above. These bands may, therefore, be of use as markers for the identification of those isolates from the individual sub-sets. (iv) Probes V1D and T1E were found to hybridise in an identical manner which is consistent with previous observations (Chapters 5 and 6). This again confirms the great similarity between these two probes. (v) RFLP analysis using V1D or T1E revealed that isolates of *R. leguminosarum* bv *trifolii* form quite a homogeneous group but that strains of *R. leguminosarum* bv *viciae* form a more diverse group. (vi) Probes V1D and T1E reveal the presence of a 2.3 Kb band in all isolates of *R. leguminosarum* bv *trifolii* and a 2.6 Kb band in most strains of *R. leguminosarum* bv *viciae*. This band size variation (about 300 bp) maybe of use for identifying isolates from these two biovars.

Strain-specific probe T37-3:

Probe T37-3, when hybridised to *R. leguminosarum* bv *trifolii* strains, produced the results shown in Figure 7.12. From observations of the results it can be seen that a band of size 2.1 Kb is strongly hybridised to in isolate *Rt*JJD15. This is the isolate from which the probe was isolated and indicates a single copy of the probe DNA is found in this strain only. The presence of this band in this isolate only indicates that this probe could be used for strain-specific research when these eighteen *R. leguminosarum* bv *trifolii* isolates are used. The absence of this probe from the seventeen isolates which fail to hybridise to it cannot be taken as indicative of its absence from a larger *R. leguminosarum* bv *trifolii* population.

The result obtained with this probe does, however, indicate that it is possible to identify and isolate strain-specific markers from RAPD profiles. However, the identification of only one probe from a potential eleven probes highlights the high level of DNA conservation within the bacterial genome.

Tables and Figures

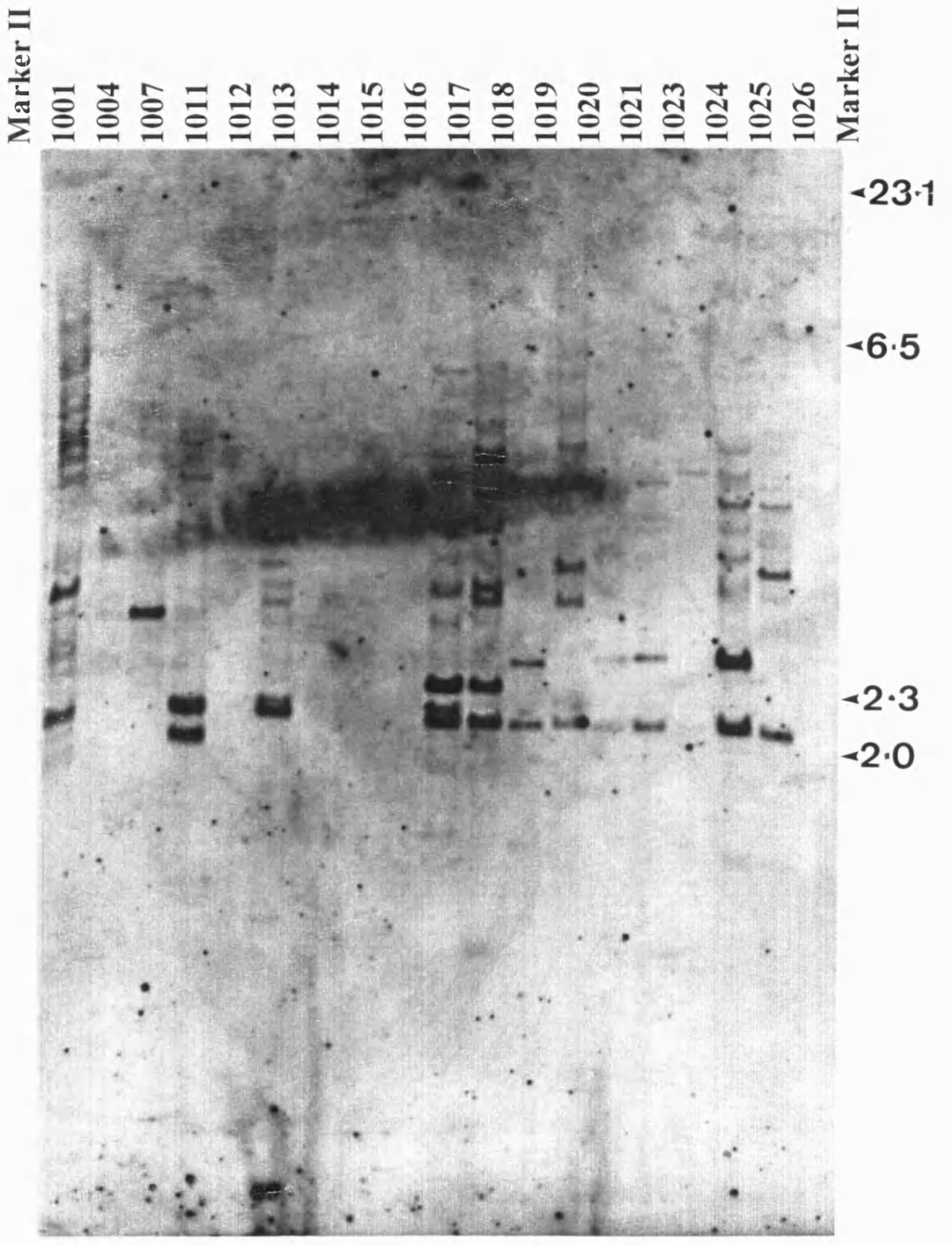


Figure 7.1: This figure shows the results of hybridising probe M1A to restriction digests of total genomic DNA from isolates of *R. leguminosarum* bv *viciae*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

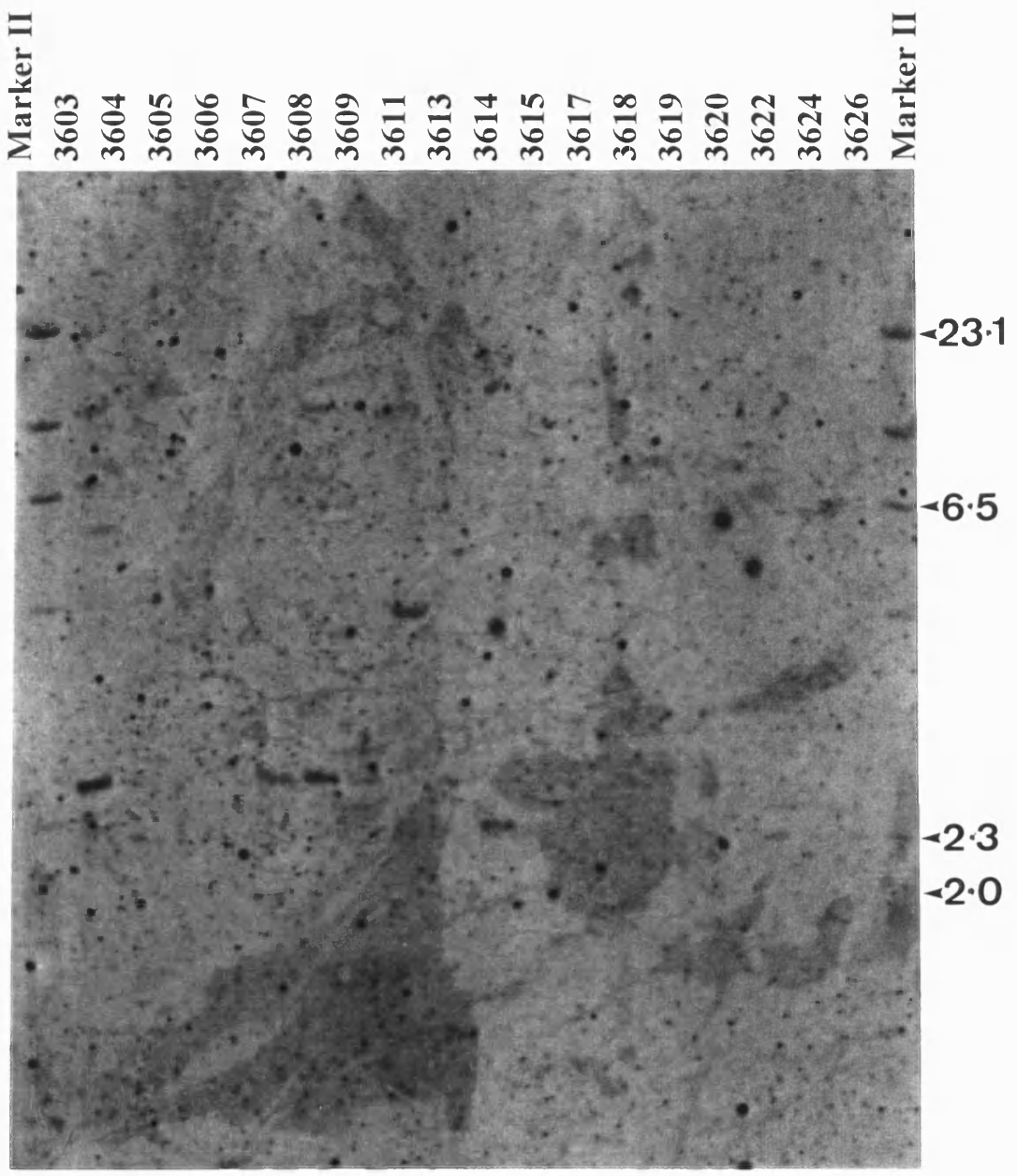


Figure 7.2: This figure shows the results of hybridising probe M1A to restriction digests of total genomic DNA from isolates of *R. leguminosarum* bv *phaseoli*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

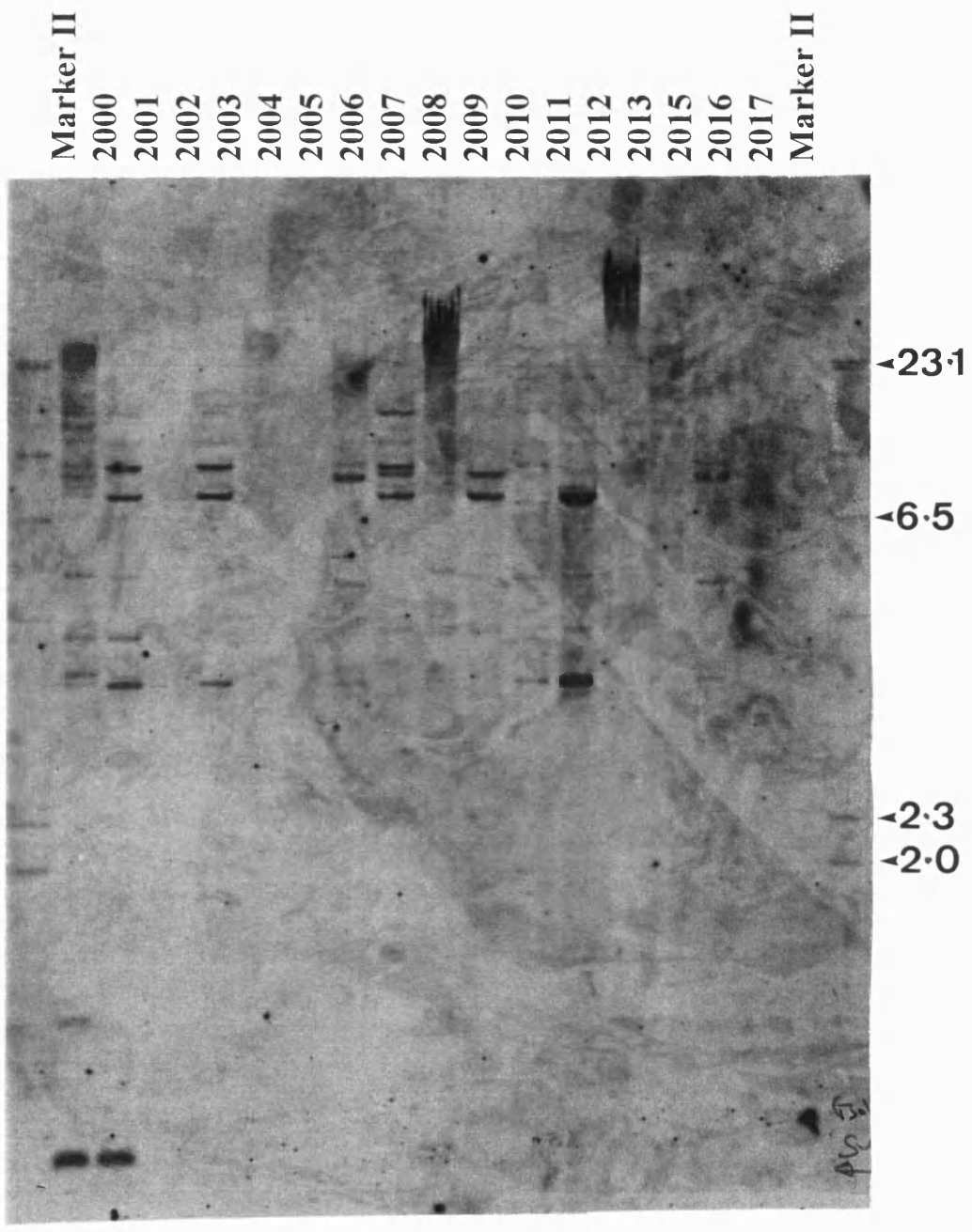


Figure 7.3: This figure shows the results of hybridising probe M1A to restriction digests of total genomic DNA from isolates of *R. meliloti*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

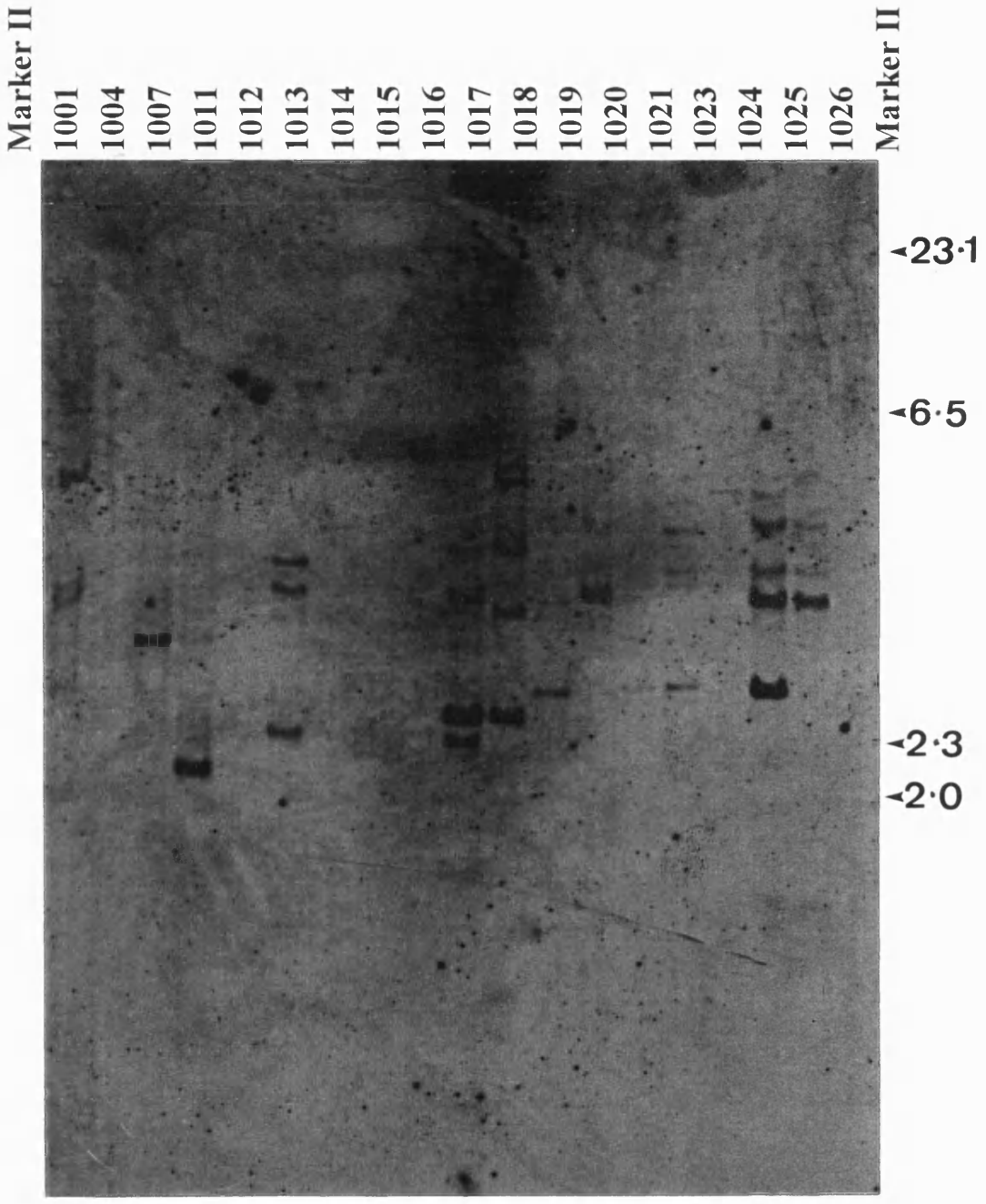


Figure 7.4: This figure shows the results of hybridising probe M1B to restriction digests of total genomic DNA from isolates of *R. leguminosarum* bv *viciae*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

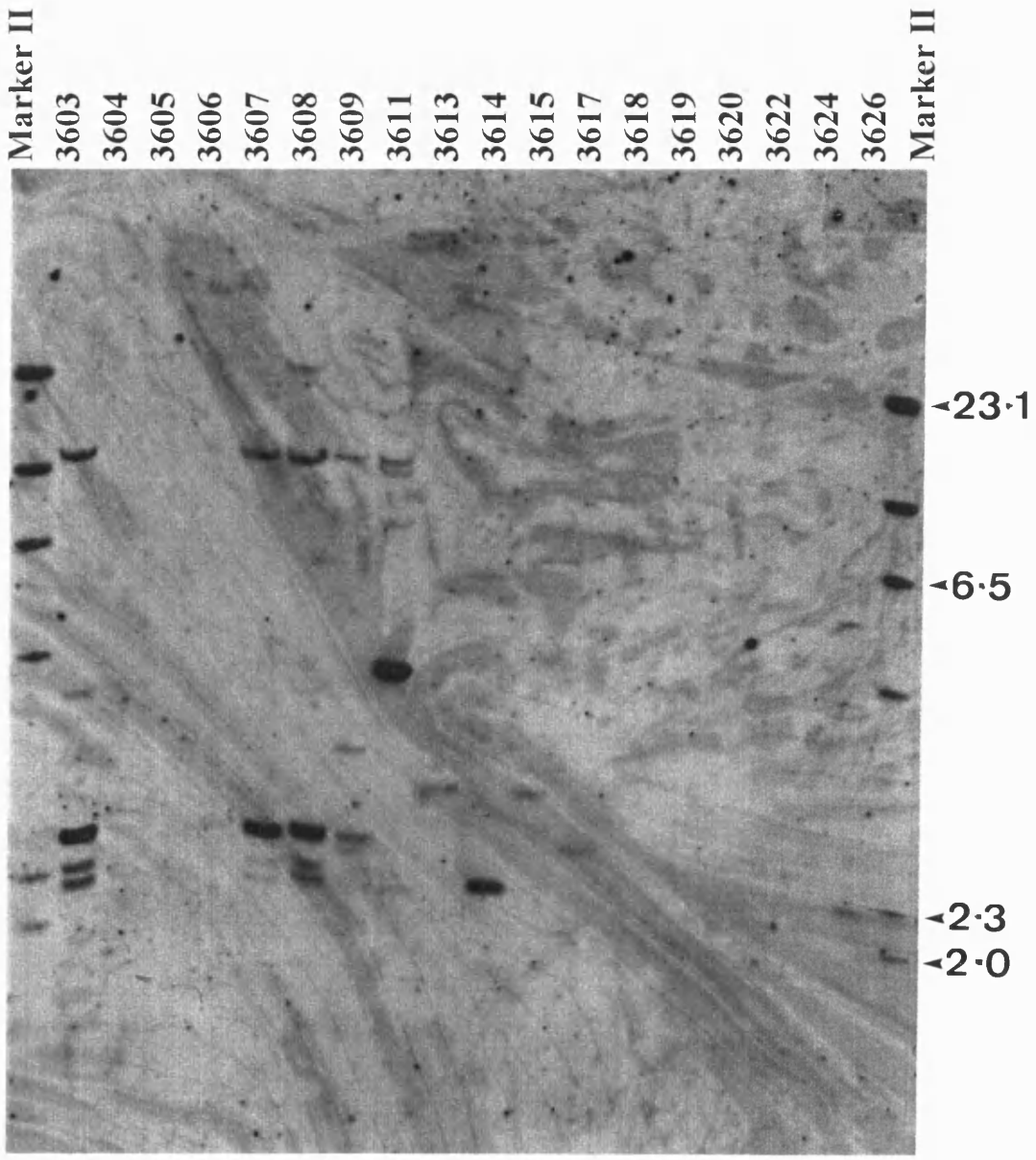


Figure 7.5: This figure shows the results of hybridising probe M1B to restriction digests of total genomic DNA from isolates of *R. leguminosarum* bv *phaseoli*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

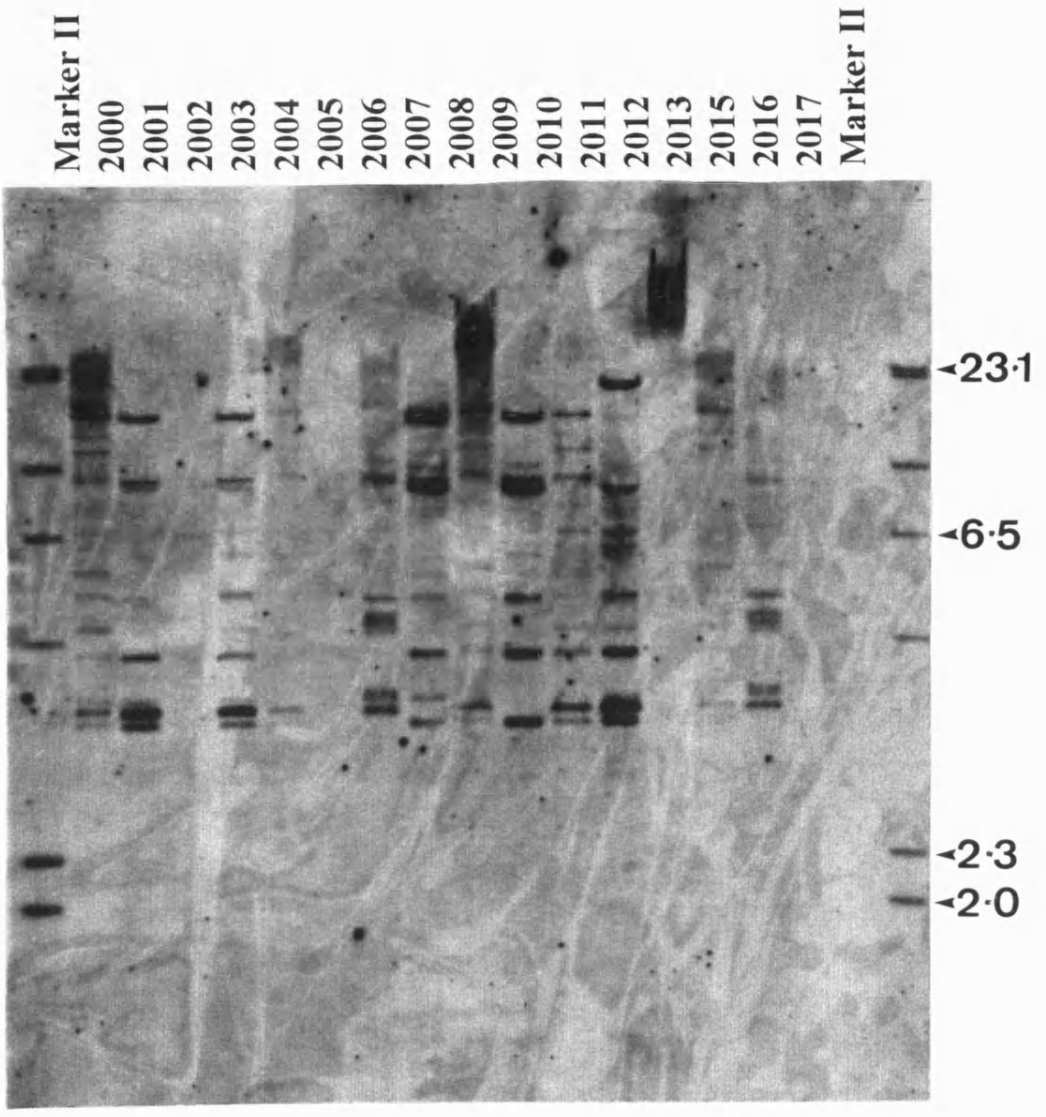


Figure 7.6: This figure shows the results of hybridising probe M1B to restriction digests of total genomic DNA from isolates of *R. meliloti*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

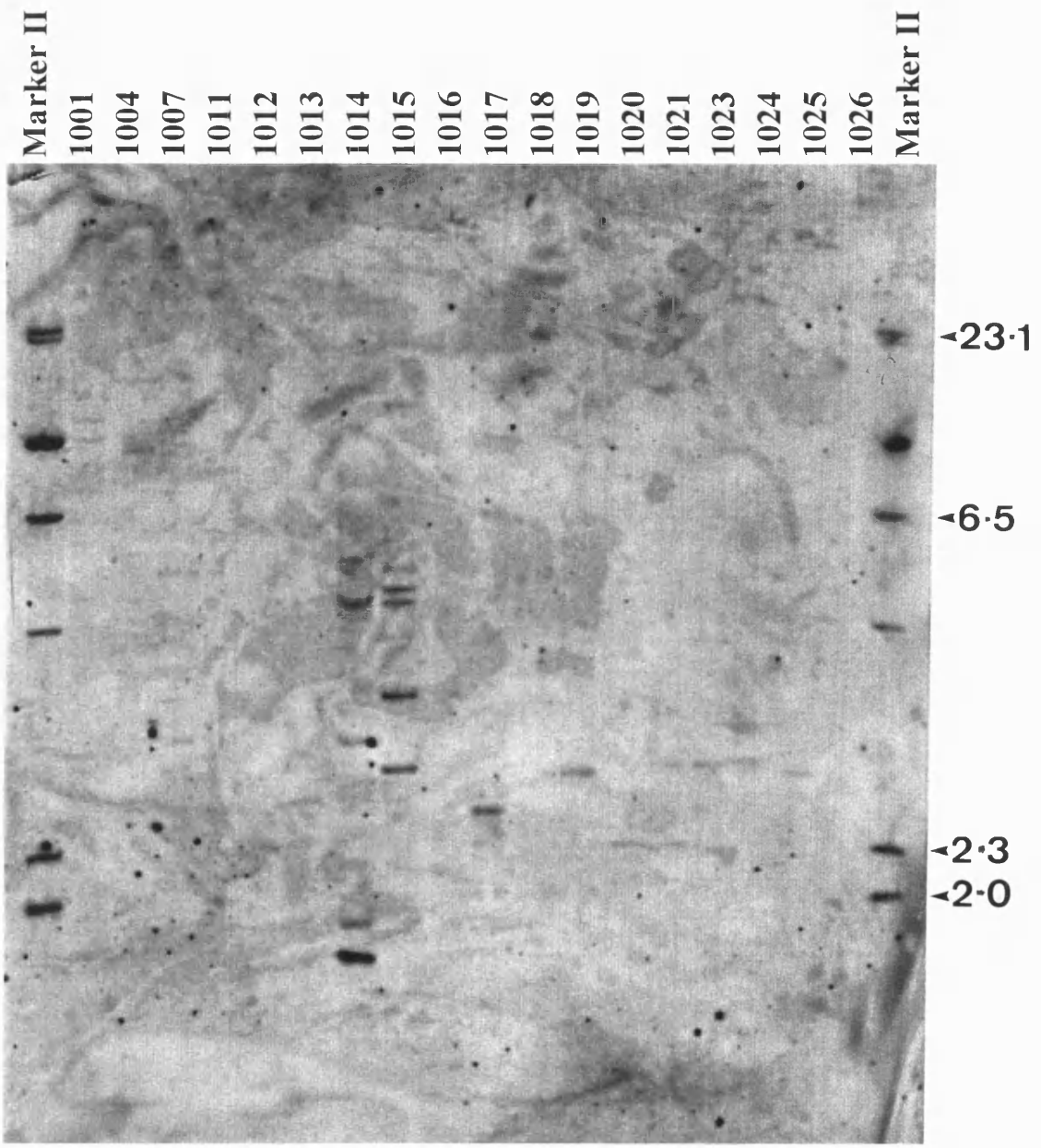


Figure 7.7: This figure shows the results of hybridising probe M1C to restriction digests of total genomic DNA from isolates of *R. leguminosarum* bv *viciae*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

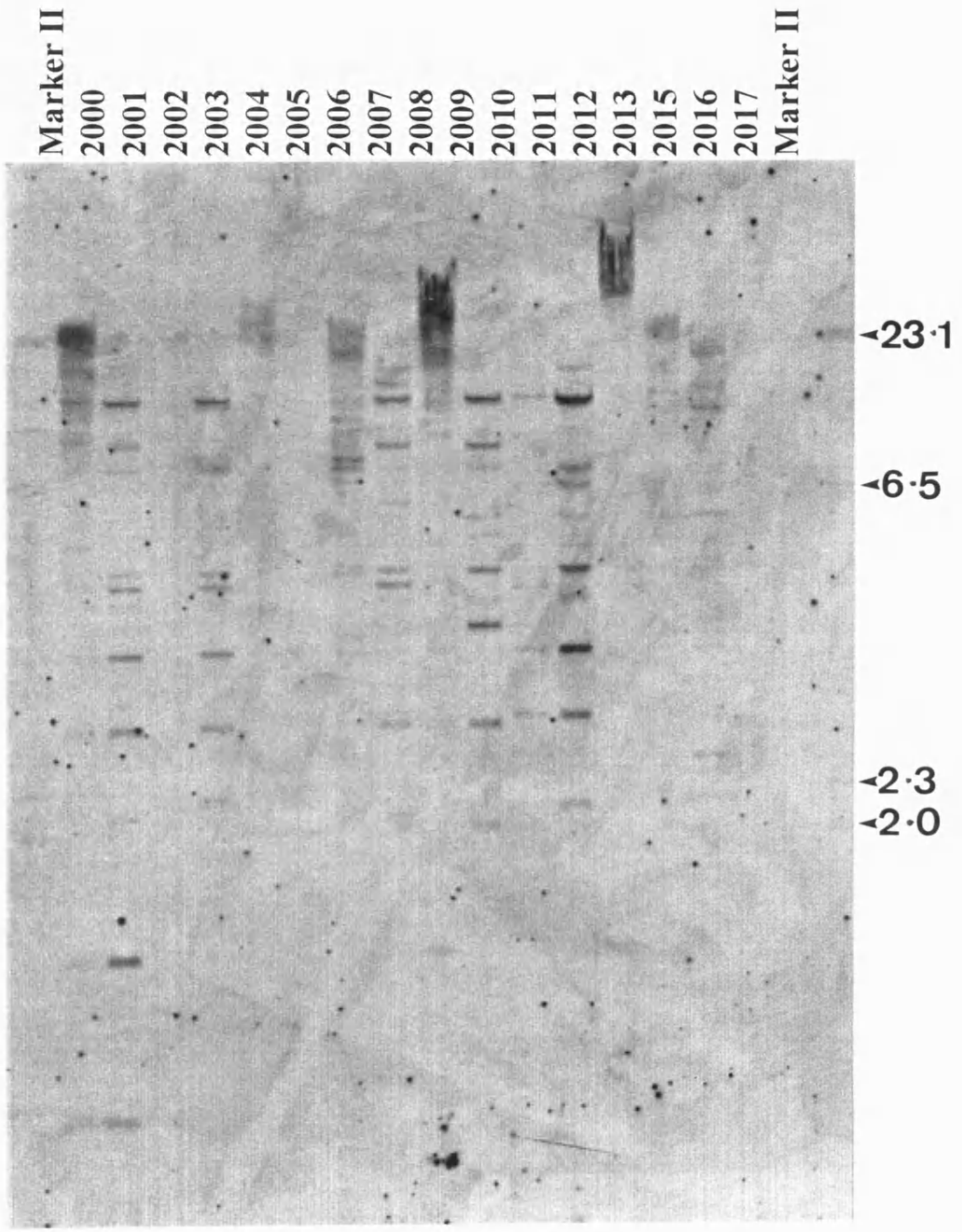


Figure 7.8: This figure shows the results of hybridising probe M1C to restriction digests of total genomic DNA from isolates of *R. meliloti*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

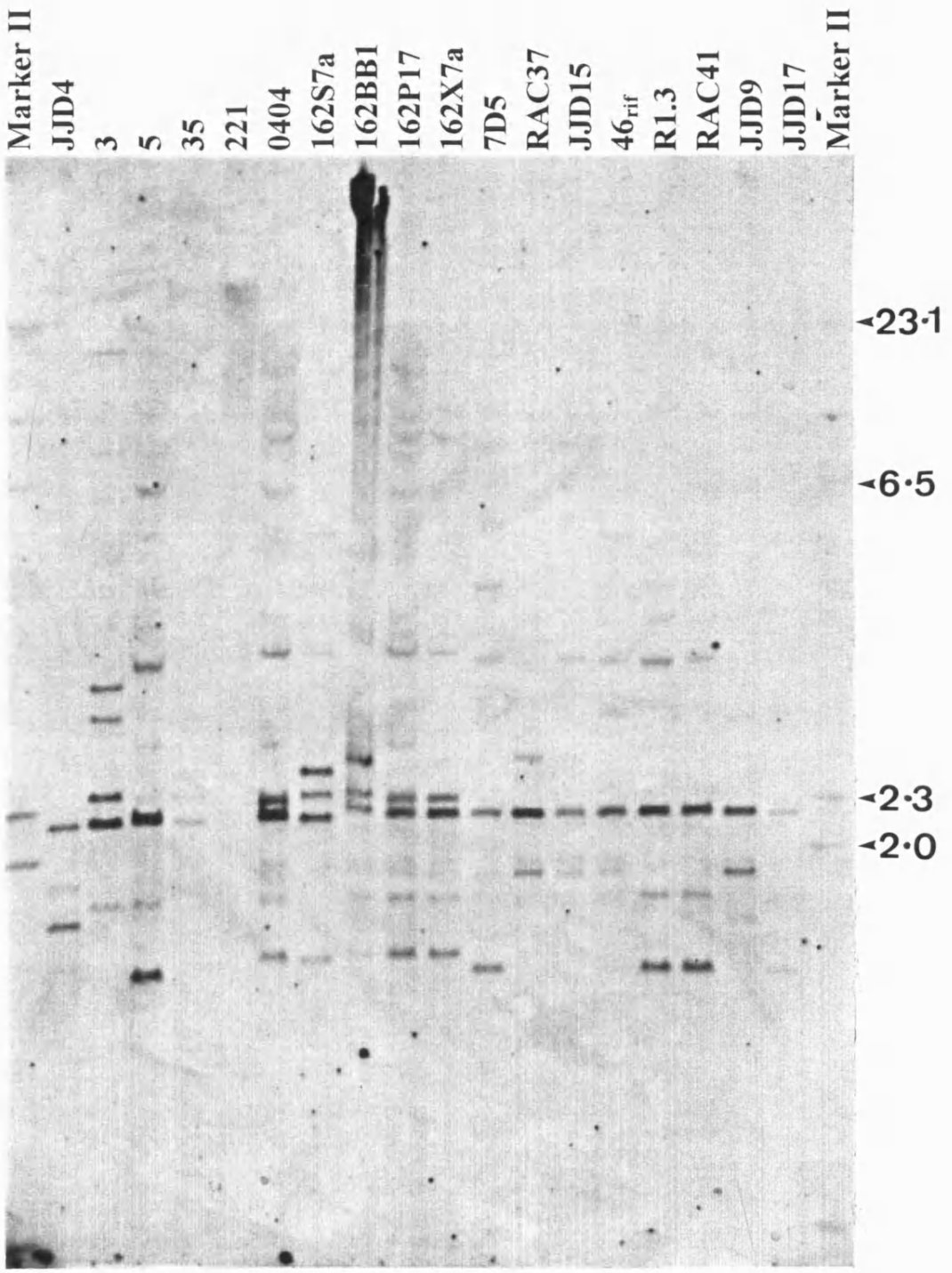


Figure 7.9: This figure shows the results of hybridising either of the probes V1D or T1E to restriction digests of total genomic DNA from isolates of *R. leguminosarum* bv *trifolii*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

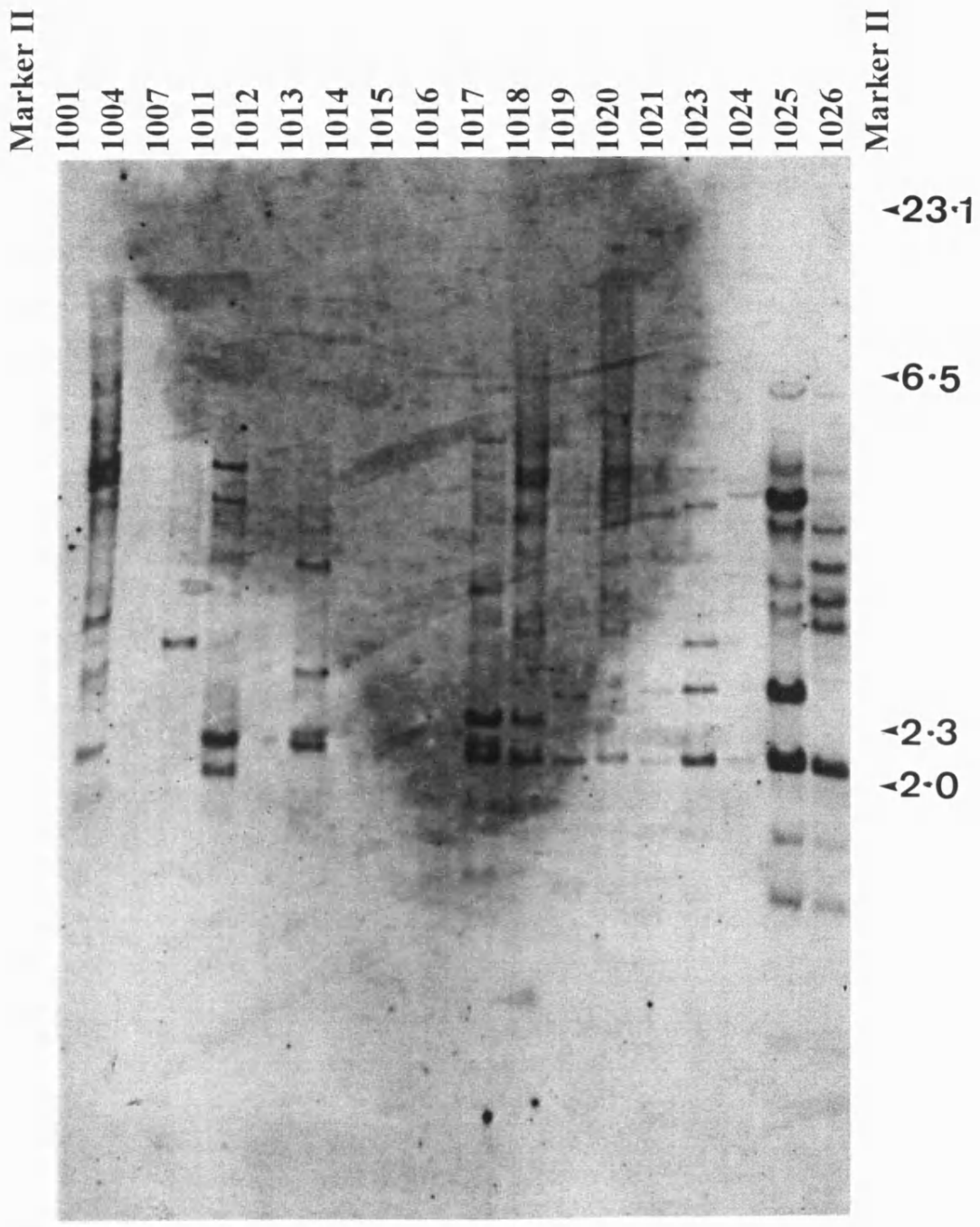


Figure 7.10: This figure shows the results of hybridising either of the probes V1D or T1E to restriction digests of total genomic DNA from isolates of *R. leguminosarum* bv *viciae*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

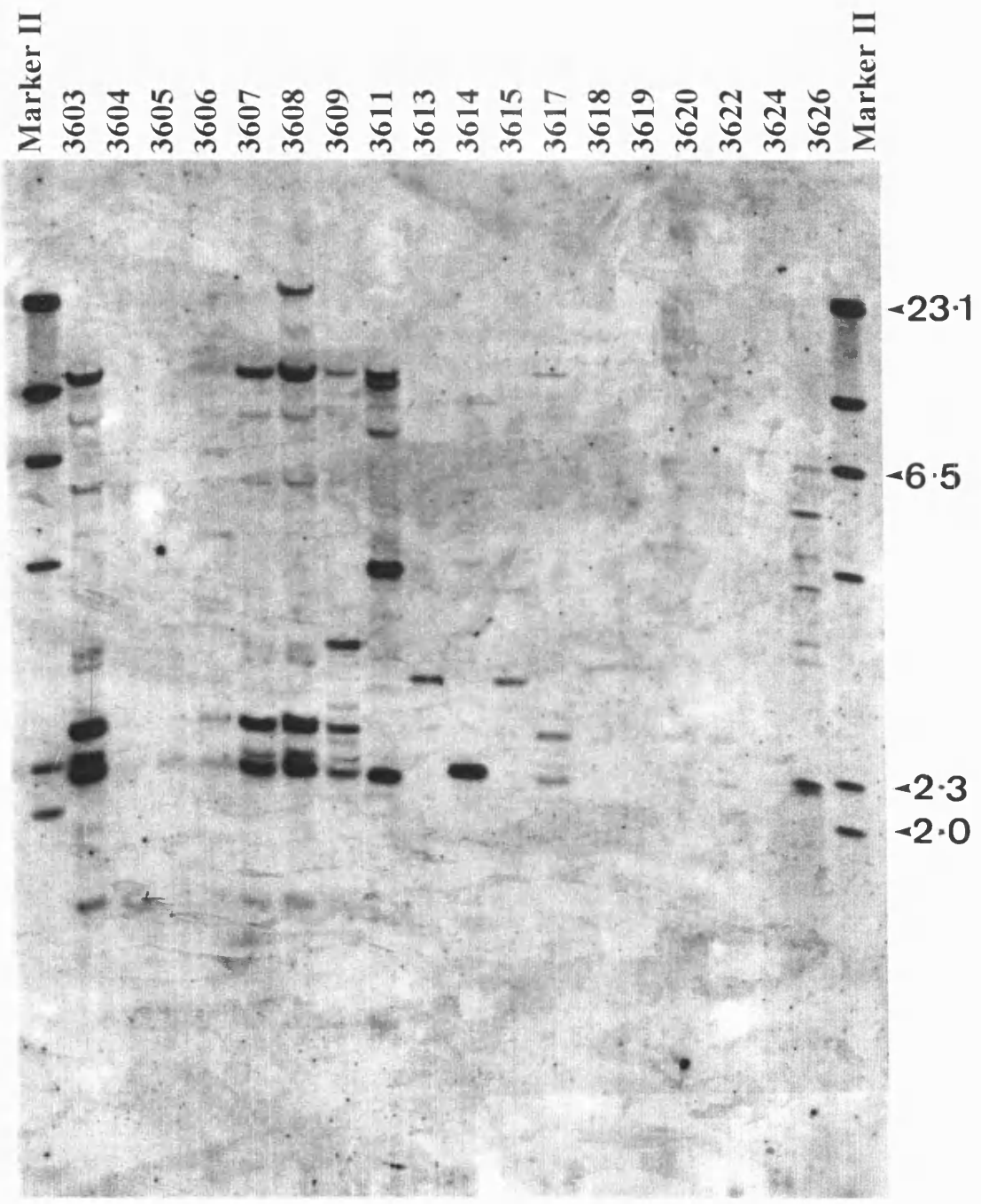


Figure 7.11: This figure shows the results of hybridising either of the probes V1D or T1E to restriction digests of total genomic DNA from isolates of *R. leguminosarum* bv *phaseoli*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

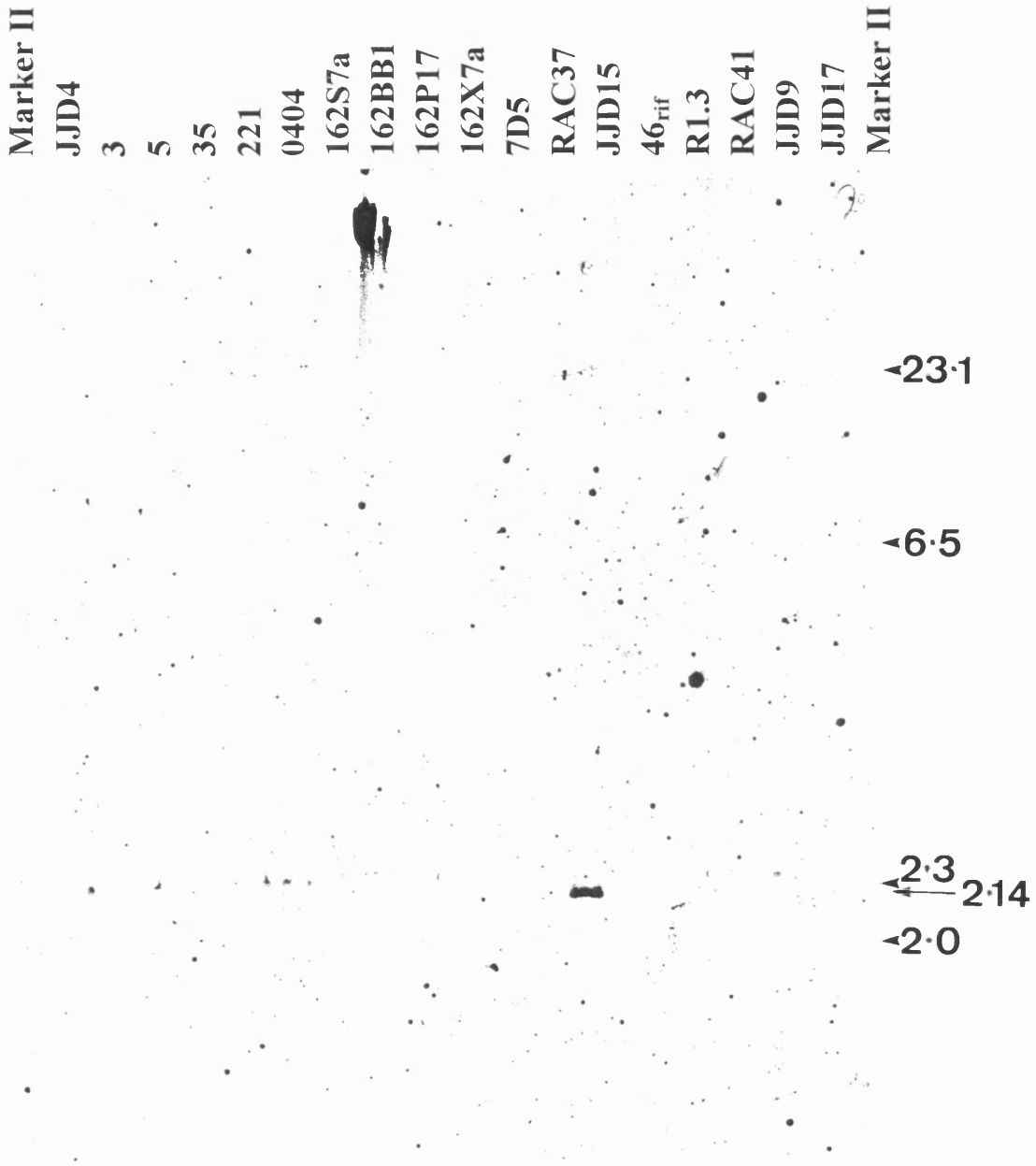


Figure 7.12: This figure shows the results of hybridising the strain-specific probe T37-3 to restriction digests of total genomic DNA from isolates of *R. leguminosarum* bv *trifolii*. Sizes (Kb) from Marker II (B. Mannheim) are shown.

General Discussion

Discussion

Rhizobium classification and taxonomy is presently in a state of flux due to the advances in molecular biology which have allowed a reassessment of the Rhizobiaceae family based on studies at the DNA level. These methods have aided with the identification of mis-labelled strains and shown that several species such as *R. meliloti*, *R. leguminosarum* bv *phaseoli* and the genera *Bradyrhizobium* are composed of at least two sub-groups (Stanley *et al.*, 1985; Young, 1985; Kuykendall *et al.*, 1988; Pinero *et al.*, 1988; Eardly *et al.*, 1990; Martínez-Romero *et al.*, 1991). At the outset of the research described in this report the most recent rhizobial taxonomy listed four genera, *Rhizobium*, *Bradyrhizobium*, *Azorhizobium* and *Sinorhizobium* with six, one, one and two recognised species respectively (Elkan, 1992). The study described in this report has been based on this review although a subsequent review by Martínez-Romero (1994) describes only three genera. The classification, from the latter review, is summarised in Table 8.1. The differences, in *Rhizobium* classification, between the two reviews did not affect the strains employed in this study, therefore the older classification was used throughout.

The aims of this research have been described previously (Chapter 1) and are reiterated below. Each objective is discussed, in turn, along with an analysis of its relevance to *Rhizobium* genetics.

Objectives:

1. To assess the use of RAPD fingerprinting for the classification and identification of *Rhizobium*, and in particular *R. leguminosarum* bv *trifolii*, isolates. This study also included the development of a suitable method of statistical analysis.

2. To identify and quantify species- and strain-specific DNA probes, arising from DNA amplification, which could be used for further studies of *Rhizobium* identification.

3. To derive genetic information regarding interspecies relationships.

4. To compare the results of these three main objectives with previously observed results.

RAPD fingerprints for Rhizobium identification and classification:

The use of RAPD fingerprints for *Rhizobium* strain identification had been described by Harrison *et al.*, (1992). Using primers from this study, in single and double primed reactions, an investigation of a larger selection of isolates from a wider number of species was undertaken.

Using the primer SPH1 (GAC_5), (which was identified from Harrison's study as potentially the best primer for identification purposes), in single primed reaction and the statistical methods of PCO and Cluster Analysis, with simple matching, it was possible to classify all the isolates into their respective species from their RAPD fingerprints (Chapter 3). This method also indicated the existence of sub-divisions within the genera *Bradyrhizobium*, the species *R. meliloti* and the biovar *R. leguminosarum* bv *phaseoli* at a level of 80-100% similarity. These divisions are in

accordance with those made by other groups (Stanley *et al.*, 1985; Young, 1985; Kuykendall *et al.*, 1988; Pinero *et al.*, 1988; Eardly *et al.*, 1990; Martínez-Romero *et al.*, 1991) and confirm the use of RAPD fingerprinting, with SPH1, as useful for *Rhizobium*, classification.

An examination of the relationships between the species, at a lower level of similarity, indicated that the species *R. meliloti* and the genera *Bradyrhizobium* are related to each other at the 75-80% level of similarity. This is evidenced by the association of the clusters from these two species. This observation tends to suggest that the *R. meliloti* is comprised of two sub-groups, but that one of these shares a degree of similarity with the lupin derived isolates of the genera *Bradyrhizobium*. This suggests that these two sub-specific groups have evolved from a common ancestor but have diverged in their nodulation host so that they are presently classified in two different genera. The observation of two sub-groups within the genera *Bradyrhizobium* has been reported previously (Stanley *et al.*, 1985; Kuykendall *et al.*, 1988) as has the detection of two sub-groups within *R. meliloti* (Young, 1985; Eardly *et al.*, 1990). However, no previous reports of an association between *Bradyrhizobium* and *R. meliloti* have been made. The results found here, therefore, tend to indicate that RAPD fingerprinting provides a different view of *Rhizobium* classification which has been over-looked by previous methods of classification. It has therefore highlighted a previously unreported relationship between *Rhizobium* and *Bradyrhizobium* and may indicate the existence of a novel species comprising strains from these two genera. However, an examination of a broader range of biochemical and molecular characteristics may verify that present classification schemes are correct but that the observations from RAPD profiles indicate an older relationship which is, today, of little importance for classification purposes. It would be interesting to examine a larger selection of strains, from these two species, in order to confirm the findings made here.

Observations of PCO plots indicate that the two biovars, *R. leguminosarum* bv *trifolii* and bv *viciae*, are very closely related but that *R. leguminosarum* bv *phaseoli* is not as related to these biovars (Figures 3.2, 4.2 and 4.5). Evidence from both spatial

orientation and Cluster Analysis, following analysis of SPH1, SPH3 and SPH3+7 primed DNA amplification reactions, supported this observation although the SPH1 data provided the clearest evidence. These findings were further supported by analysis of RAPD profiles from *R. leguminosarum* strains only (Figures 4.6 and 4.7) where it can be seen that *R. leguminosarum* bv *trifolii* and bv *viciae* isolates form a straight line, on the plot, whereas isolates of *R. leguminosarum* bv *phaseoli* form a scattered group. Pinero et al., (1988) have also indicated that *R. leguminosarum* bv *phaseoli* is a mixed group of up to seven distinct strain types. The PCO plot produced after analysis of SPH1 profiles, however, indicates that *R. leguminosarum* bv *phaseoli* may be related to the other *R. leguminosarum* isolates to a small degree. This may be due to a common ancestor which is related to strains from either *R. leguminosarum* bv *trifolii* or bv *viciae* and to one of the sub-groups forming the cluster *R. leguminosarum* bv *phaseoli*.

The use of two primers to produce RAPD fingerprints has been reported (Welsh & McClelland, 1991; Caetano-Anolles *et al.*, 1991; Fekete *et al.*, 1992). This method was also used to investigate the *Rhizobium* isolates studied previously with primer SPH1. The PCO ordination plot produced from results obtained using the primers SPH3 (GACGACAGCGGC) and SPH7 (CAGCCACAGCGC) in a double primed reaction shows species separation similar to that observed with SPH1, with the exception that the *R. meliloti* and *Bradyrhizobium* are somewhat intermingled spatially (Chapter 4). This observation again appears to support the hypothesis that *R. meliloti* and *Bradyrhizobium* are related to a certain degree. This ordination plot also resulted in a better resolution of the *R. leguminosarum* bv *trifolii* and bv *viciae* isolates which were indistinguishable from analysis of SPH1 derived RAPD fingerprints. However, upon application of Cluster Analysis this apparent separation was found to be incorrect. This result highlighted one of the problems of taking ordination plots at face value, i.e. spatial distribution on the plot cannot be used for absolute relatedness purposes, but only as a guide to the relative relatedness of isolates. In this case, however, it may be possible to use the spatial distribution to classify isolates into the

two biovars (bv *trifolii* and bv *viciae*) but not to show absolute relatedness amongst the strains. The spatial orientation observed does, however, indicate that detectable differences exist between the two biovars and that with the use of the correct primers it should be possible to differentiate between them.

General observations indicated that the use of primers SPH3 and SPH7, in a double primed reaction, is of little use for *Rhizobium* classification. However, the results obtained do indicate that double primed RAPD profiles may be of use for *Rhizobium* classification purposes although further investigations would be necessary.

As a final assessment of the RAPD fingerprints, data from different amplification reactions was grouped and the new set analysed using PCO and Cluster Analysis with simple matching. Such grouped data was strongly influenced by the results from SPH1 and SPH3 primed reactions. The SPH1 data acted in a positive manner whilst the SPH3 results were detrimental to the overall analysis. Combined data from SPH1 and SPH3+7 primed reactions appeared to be the most useful, providing a relatively accurate way to separating *R. leguminosarum* bv *trifolii* isolates from *R. leguminosarum* bv *viciae* isolates. However, in general combining data from several amplification reactions was not useful for classifying *Rhizobium* strains.

Statistical analysis:

When this study was initiated little work on the use of RAPD fingerprints for classification purposes had been reported. There was also a scarcity of statistical analysis methods which could be employed to analyse RAPDs from a large strain selection such as that used here. What was needed was a technique which could assess the relatedness of the strains based on the presence (1) or absence (0) of RAPD bands. An assessment based on the relative intensity of the individual bands was inappropriate due to fluctuations in band intensity observed between reactions. The statistical method employed also needed to clearly show trends in the data so that an assessment of the use of RAPDs for classification could be made, i.e. it was necessary to determine if it was even possible to classify *Rhizobium* strains using RAPD profiles.

The use of Principal Coordinates Analysis (PCO) to analyse presence/absence (1/0) data, as reported by Digby and Kempton (1987) appeared to meet the criteria needed to examine RAPD fingerprints.

Using this method it was possible to determine strain groups based on spatial orientation on the PCO plot. The additional application of Cluster Analysis, using the simple matching coefficient, confirmed the existence of these groups and also indicated the existence of sub-groups within the larger clusters.

The use of the simple matching coefficient, for Cluster Analysis, was confirmed by repeating the analysis using the Jaccard matching coefficient. An assessment of the RAPD data using the Jaccard matching coefficient, although useful for species identification, failed to indicate the sub-species divisions within certain species. This observation indicated that Jaccard matching was inappropriate for RAPD profile analysis. The inferior results obtained with the Jaccard matching method were probably due to the nature of the technique which eliminates those loci which reveal no band in either strain. This differs from the simple matching method which would regard these loci as 100% similar. The simple matching technique is a better method of analysing RAPD fingerprints as each band is the product of two independent loci, the primer binding sites. This is taken into account slightly more with the simple matching method than with the Jaccard matching method.

The observations made from analysis of RAPD fingerprints with PCO and Cluster Analysis indicate that it is possible to use RAPD fingerprints to classify *Rhizobium* strains. However, PCO analysis should only be used as a guide to highlight trends in data therefore may not be the most appropriate method for analysing RAPD data. Recent reports of the use of the statistical techniques of AMOVA and HOMOVA to study RAPD profiles have been made (Haig *et al.*, 1994; Stewart & Excoffier, 1996; Stewart *et al.*, 1996). It would be interesting to compare the use of these methods, to analyse the RAPD profiles described here, with the PCO and Cluster Analysis used in this study.

The results from this study indicate that it is possible to employ RAPD fingerprinting techniques to classify strains of *Rhizobium* and *Bradyrhizobium*. However, with the primers in this study (SPH1, SPH3 and SPH3+7) it is not possible to categorically identify strains of *R. leguminosarum* bv *trifolii*. Results from RAPDs produced using primers SPH3+7 do, however, indicate that discernible differences exist between *R. leguminosarum* bvs *trifolii* and *viciae* so that with the use of the correct primers it should be possible to clearly classify or identify isolates of these two biovars. The use of other primers may also permit individual *Rhizobium* strain identification so that they can be studied in the natural environment. Until then the use of DNA probes provides one method of investigating *Rhizobium* interactions in soil samples.

RAPD-derived DNA probes:

Species-specific:

Evidence of bands common to the majority of isolates, within a single species, implies that these may be species-specific and therefore of potential use as species-specific DNA probes. RAPD products such as these were identified in *R. meliloti* and *R. leguminosarum* bvs *trifolii* and *viciae* and were isolated using the technique of band stabbing (Bjourson and Cooper, 1992). This provided a rapid method of producing DNA probes without the need to clone fragments.

These probes were screened for their suitability as species-specific probes by hybridising them to Southern blots of RAPD profiles (Chapter 5), restriction digests of total genomic DNA (Chapter 7) and colony blots (Chapter 6). The latter two methods were found to be more discriminating when assessing the specificity of the probes. This is probably due to the nature of the RAPD profile, which only comprises a small percentage of the total bacterial genome therefore can easily be missing DNA which is homologous to the probes. However, results of back-hybridisation to the RAPD

profiles did indicate that probes V1D and T1E, although varying in size by 40 bp, were probably from similar regions of their respective genomes. Back-hybridising to the RAPDs also suggested that probes M1A and M1B share a degree of homologous DNA but that both show no homology with probe M1C.

The results from colony and RFLP analysis (Chapters 6 and 7) again support the theory that probes V1D and T1E are similar as both reveal identical hybridisation patterns from both tests. Results from colony blots suggest that these probes may not be species-specific as they show a degree of homology with some strains of *R. meliloti* and may even share homologous DNA with *Bradyrhizobium* isolates. This observation was, however, not confirmed when these probes were hybridised to total genomic DNA digests. This anomaly may be the result of DNA fixing to cellular debris on the colony blots as the signal was quite weak and there was only one isolate (*Rm*2005) which indicated any homology with the probes.

Cross-hybridisational activity was revealed when probes M1A, M1B and M1C, were screened against colony blots. These probes were seen to hybridise with strains of *Bradyrhizobium* and *R. leguminosarum* which indicated they were of little use for species-specific studies. Probe M1A was found to hybridise to four *Bradyrhizobium* isolates which had previously been found to cluster with the *R. meliloti* group following analysis of RAPD fingerprints. These two observations suggest a degree of conservation exists between these strains. This result may, however, be in support of those made previously by Stanley *et al.*, (1985) and Kuykendall *et al.*, (1988) who both suggested that *Bradyrhizobium* may be a mixture of several strain types. Sadowsky *et al.*, (1987a) have suggested that the species *R. fredii* (now *S. fredii* [Chen *et al.*, 1988]) may be an evolutionary link between *Bradyrhizobium* and *R. meliloti*. The strains used in this study have not been studied further, therefore, it is unknown if they should be reclassified as *S. fredii*. However, the results obtained here do indicate that their present classification is inaccurate. This observation is supported by colony blot results from M1C which reveal hybridisation between the four *Bradyrhizobium* isolates and this *R. meliloti* probe. Both probes, M1A and M1C, also

hybridise, to a degree, with *R. leguminosarum* isolates which indicates the existence of a relationship between the species *R. meliloti* and *R. leguminosarum*. This could be accounted for by plasmid transfer which has been previously reported to occur between these two species (Djordjevic *et al.*, 1983; Broughton *et al.*, 1987). The low level of cross-reactivity observed between these two species is in accordance with an infrequent transfer of genetic material between the species and indicates that they are not closely related. It is also possible that the strains showing cross-hybridisational activity with the *R. meliloti* probes are of the species *R. tropici* (Martínez-Romero *et al.*, 1991). This species comprises isolates formerly classified as *R. leguminosarum* bv *phaseoli* Type II, however, there is presently insufficient information about the strains used in this study to allow reclassification into this new species. The observations made here could, therefore, also indicate a possible relationship between the species *R. tropici* and *R. meliloti*.

RFLP analysis intimates that the strains comprising *R. leguminosarum* bv *trifolii* are very homogeneous whilst those from *R. leguminosarum* bv *viciae* are more diverse. Two bands of 2.3 Kb and 2.6 Kb were found in *R. leguminosarum* bv *trifolii* and *R. leguminosarum* bv *viciae* respectively. These reveal that variations do occur between the two biovars and suggest that the two different sized bands may be of use for differentiating amongst strains of *R. leguminosarum* bvs *trifolii* and *viciae*. Within the species *R. meliloti* similar markers of 13.8 Kb, 5.3 Kb, 4.2 Kb, 3.5 Kb and 3.3 Kb were found when RFLP patterns were produced using M1B. The presence or absence of these five bands from the *R. meliloti* isolates corresponds to *R. meliloti* groupings identified from RAPD profile analysis (Chapter 3). These markers, again, should be useful for strain differentiation. An analysis of overall RFLP patterns, produced with M1B, resulted in an identical strain classification as that obtained from RAPD profile analysis. These observations, therefore, lend further support to the previous findings of sub-divisions within the species *R. meliloti* (Young, 1985; Eardly *et al.*, 1990; Dooley *et al.*, 1993) and those described in this report. However, results from hybridising M1A to restriction digests indicates that both sub-groups of *R. meliloti* share a degree

of homologous DNA. This, therefore, suggests that although variations do occur in this species the two major clusters are related. This again supports the theory that one of the *R. meliloti* clusters forms an evolutionary link between *R. meliloti* and *Bradyrhizobium*. .

In general, none of the potential species-specific probes, M1A, M1B, MIC, V1D and T1E, were found to be totally specific, and M1B even appeared to be universally distributed. Some of the probes, however, may be of use in controlled environments were a limited number of isolates are used. The probes of greatest value were M1B and V1D / T1E which provided a potential method for the identification and classification of strains from their respective species, *R. meliloti* and *R. leguminosarum* respectively.

Strain-specific:

Following analysis of colony blots (Chapter 6) only one potential strain-specific probe, T37-3, appeared to be truly strain-specific. Several of the other potential strain-specific probes appeared useful if used in experiments with a limited number of isolates or as species-specific probes. Only probe T37-3 was, therefore, used to hybridise Southern blots of total genomic DNA digests (Chapter 7). This low number of true strain-specific probes indicates the close relationship between isolates of the same species and intimates at the scarcity of superfluous DNA which would be of use to produce strain-specific probes. The presence of apparent strain-specific bands in the RAPD fingerprints of some isolates only indicates that there is some DNA variation between the strains. This is apparent by the emergence of one, or more, of the primer binding sites, in individual isolates, which has resulted in the appearance of potential strain-specific bands.

Analysis of restriction digested DNA with the probe T37-3 revealed the presence of a single band of 2.1 Kb in its derivative strain, *Rt*JJD15. This indicated it is of use as a strain-specific probe. It should be noted, however, that the strain-specificity

of this probe is in relation to the eighteen isolates employed in this study and cannot be interpreted as a global specificity.

This result does, however, verify that RAPD profiles can be utilised to produce strain-specific DNA probes.

An assessment of interspecies relationships:

The present taxonomic classification of *Rhizobium* species has been summarised in Table 8.1. The results obtained from this study have mirrored this classification system in general although some deviation has been observed. An incomplete taxonomic examination of the strains comprising *R. leguminosarum* bv *phaseoli*, *R. meliloti* and *Bradyrhizobium*, as described in this report, has meant it is not possible to draw absolute conclusions about interspecies relationships. This has meant it was not possible to determine if some of the strains employed, in this work, should be assigned to the new groups of *R. tropici*, *R. fredii* or *B. elkanii*. However, the analysis of RAPD fingerprints, especially those produced with the primer SPH1, has shown that the *Rhizobium* species are related at varying degrees of similarity. Most interesting of these relationships was that between *R. meliloti* and *Bradyrhizobium* which emerged at the 75-80% level of similarity. It was also determined that the biovars *trifolii* and *viciae* of the species *R. leguminosarum* are not as closely related to the other biovar (bv *phaseoli*) as they are to each other. It was also obvious that within biovar *phaseoli* there were smaller sub-groups which displayed varying degrees of relatedness to each other. This was in accordance with previously reports (Pinero *et al.*, 1988; Martínez-Romero *et al.*, 1991). Some of these interspecies relationships were confirmed by the use of DNA probes, derived from the RAPD fingerprints. It was found that the *R. meliloti* probes hybridised to those *Bradyrhizobium* strains which clustered with the *R. meliloti* following RAPD fingerprint analysis. The two *R. leguminosarum* probes also reacted in a similar way even though they came from the two biovars, *R. leguminosarum* bv *trifolii* and bv *viciae*. This highlighted the close

relationship which exists between these two biovars. The use of the *R. meliloti* probe, M1B, revealed that the sub-groups within the species *R. meliloti*, which had been observed following RAPD analysis, were detectable from RFLP profiles. This method also indicated that there may be a second sub-group within one of the sub-groups, which suggests that *R. meliloti* is comprised of three groups of related strains, two of which are highly related. These highly related clusters are not detectable when RAPDs are produced with the primer SPH1.

In conclusion, the use of molecular techniques for the study of *Rhizobium* genetics is aiding the general understanding of their taxonomic groupings and classification. This has led to changes in the old species definitions which will no doubt continue for sometime to come. The report by Martínez-Romero, (1994), which contains a description of the most recent taxonomic classification has been summarised in Table 8.1.

Many of the molecular techniques used to classify bacterial isolates involve an examination of small, specific parts of the genome which is useful for individual strain identification but may not be appropriate for taxonomic classification. The use of fingerprinting techniques, especially RAPDs, examines variations over the whole genome which, as this report has shown, is useful for species and strain identification. This report has highlighted the use of RAPD fingerprinting for confirming previous observations and for indicating possible unnoticed relationships within the Rhizobiaceae.

Tables and Figures

Table 8.1: Taxonomy of the Rhizobiaceae family (from Martínez-Romero, 1994).

Recognised genera	Recognised species
Rhizobium	
	R. etli
	R. fredii
	R. galegae
	R. huakuii
	R. leguminosarum
	bv *viciae*
	bv *trifolii*
	bv *phaseoli*
	R. loti
	R. meliloti
	R. tropici
	R. xinjiangensis
Bradyrhizobium	
	B. elkanii
	B. japonicum
Azorhizobium	
	A. caulinodans

References

References

AHMAD M.H., EAGLESHAM A.R.J. & HASSOUNA S. (1981). Examining serological diversity of "cowpea" *Rhizobia* by the ELISA technique. Archives of Microbiology. **130**, 281-287.

ALLEN O.N. & BALDWIN I.L (1931). The effectiveness of rhizobia. Wisconsin Agricultural Experimental Station Research Bulletin, number 106.

AMARGER N., BOURS M., REVOY F., ALLARD M.R. & LAGUERRE G. (1994). *Rhizobium tropici* nodulates field-grown *Phaseolus vulgaris* in France. Plant and Soil. **161**, 147-156.

AMES-GOTTFRED N.P., CHRISTIE B.R. & JORDAN D.C. (1989). Use of the Chrome Azurol S agar plate technique to differentiate strains and field isolates of *Rhizobium leguminosarum* biovar *trifolii*. Applied and Environmental Microbiology. **55**, 707-710.

AYALA F.J. (1982). Genetic variation in natural populations: Problems of electrophoretically cryptic alleles. Proceedings of the National Academy of Sciences. **79**, 550-554.

BAILEY E. & LEAR T.L. (1994). Comparison of Thoroughbred and Arabian horses using RAPD markers. Animal Genetics. **25**, 105-108.

BERESWILL S., PAHL A., BELLEMANN P., ZELLER W. & GEIDER K. (1992). Sensitive and species-specific detection of *Erwinia amylovora* by polymerase chain reaction analysis. Applied and Environmental Microbiology. **58**, 3522-3526.

BERGER J.A., MAY S.N., BERGER L.R. & BOHLOOL B.B. (1979). Colorimetric enzyme-linked immunosorbent assay for the identification of strains of *Rhizobium* in culture and in the nodules of lentils. Applied and Environmental Microbiology. **37**, 642-646.

BERINGER J.E. (1974). R factor transfer in *Rhizobium leguminosarum*. Journal of General Microbiology. **84**, 188-198.

BIETZ J.A. & SIMPSON D.G. (1992). Electrophoresis and chromatography of wheat proteins: available methods and procedures for statistical evaluation of the data. Journal of Chromatography. **624**, 53-80.

BJOURSON A.J. & COOPER J.E. (1988). Isolation of *Rhizobium loti* strain-specific DNA sequences by subtraction hybridization. Applied and Environmental Microbiology. **54**, 2852-2855.

BJOURSON A.J., STONE C.E. & COOPER J.E. (1992). Combined subtraction hybridisation and polymerase chain reaction amplification procedure for isolation of strain-specific *Rhizobium* DNA sequences. Applied and Environmental Microbiology. **58**, 2296-2301.

BJOURSON A.J. & COOPER J.E. (1992). Band-stab PCR: a simple technique for the purification of individual PCR products. Nucleic Acid Research. **20**, 4675.

BOSCO M., LUMINI E. & NORMAND P. (1996). PCR-RFLP direct fingerprinting of uncultured *Frankia* microsymbiont from *Dryas drummondii* nodules. Annali di Microbiologia ed Enzimologia. **46**, 115-123.

BOTSTEIN D., WHITE R.L, SKOLNICK M. & DAVIS R.W. (1980). Construction of a genetic linkage map in man using Restriction Fragment Length Polymorphisms. American Journal of Human Genetics. **32**, 314-331.

BROCKMAN F.J. & BEZDICEK D.F. (1989). Diversity within serogroups of *Rhizobium leguminosarum* biovar *viceae* in the Palouse region of eastern Washington as indicated by plasmid profiles, intrinsic antibiotic resistance and topography. Applied and Environmental Microbiology. **55**, 109-115.

BROMFIELD E.S.P., STEIN M. & WHITE R.P. (1982). Identification of *Rhizobium* strains on antibiotic concentration gradients. Annals of Applied Biology. **101**, 269-277.

BROMFIELD E.S.P., THURMAN N.P., WHITWELL S.T. & BARRAN L.R. (1987). Plasmids and symbiotic effectiveness of representative phage types from two indigenous populations of *Rhizobium meliloti*. Journal of General Microbiology. **133**, 3457-3466.

BROUGHTON W.J., SAMREY U. & STANLEY J. (1987). Ecological genetics of *Rhizobium meliloti*: symbiotic plasmid transfer in the *Medicago sativa* rhizosphere. FEMS Microbiology Letters. **40**, 251-255.

BRUCE K.D., HIORNS W.D., HOBMAN J.L., OSBORN A.M., STRIKE P. & RITCHIE D.A. (1992). Amplification of DNA from native populations of soil bacteria by using the polymerase chain reaction. Applied and Environmental Microbiology. **58**, 3413-3416.

BRUNEL B., ROME S., ZIANI R. & CLEYETMAREL J.C. (1996). Comparison of nucleotide diversity and symbiotic properties of *Rhizobium meliloti* populations from annual Medicago species. FEMS Microbiology Ecology. **19**, 71-82.

BURKARDT B. & BURKARDT H-J. (1984). Visualization and exact molecular weight determination of a *Rhizobium meliloti* megaplasmid. Journal of Molecular Biology. **175**, 213-218.

BURR T.J., REID C.L., YOSHIMURA M., MOMOL E.A. & BAZZI C. (1995). Survival and tumorigenicity of *Agrobacterium vitis* in living and decaying grape roots and canes in soil. Plant Disease. **79**, 677-682.

CAETANO-ANOLLÉS G., BASSAM B.J. & GRESSHOFF P.M. (1991). DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Biotechnology. **9**, 553-557.

CAMPBELL R. (1985). Plant microbiology. Edward Arnold (Publishers) Ltd., London, Victoria, Baltimore.

CARRASCO A., SALYARDS J.R. & BERRY A.M. (1995). Studies of 2 *Frankia* strains isolated from *Trevoa trinervis* Miers. Plant and Soil. **171**, 359-363.

CASSE F., BOUCHER C., JULLIOT J.S., MICHEL M. & DENAIRE J. (1979). Identification and characterisation of large plasmids in *Rhizobium meliloti* using agarose gel electrophoresis. Journal of General Microbiology. **113**, 229-242.

CASTAGNA R., MAGA G., PERENZIN M., HEUN M. & SALAMINI F. (1994). RFLP based genetic-relationships of einkorn wheats. Theoretical And Applied Genetics. **88**, 818- 823.

CHAKRABORTI R. & RYMAN N. (1983). Relationship of mean and variance of genotypic values with heterozygosity per individual in a natural-population. Genetics. **103**, 149-152.

CHANWAY C.P. & HOLL F.B. (1986). Suitability of intrinsic antibiotic resistance as a method of strain identification in *Rhizobium trifolii*. Plant and Soil. **93**, 287-291.

CHEN W.X., YAN G.H & LI J.L. (1988). Numerical taxonomic study of fast-growing soybean rhizobia and a proposal that *Rhizobium fredii* be assigned to *Sinorhizobium* gen. nov. International Journal of Systematic Bacteriology. **38**, 392-397.

CHEN W.X., WANG E., WANG S.Y., LI Y.B., CHEN X.Q. & LI Y.B. (1995). Characterisation of *Rhizobium tianshanense* Sp Nov, a moderately and slowly growing root nodule bacterium isolated from an arid saline environment in Xinjiang, Peoples Republic of China. International Journal of Systematic Bacteriology. **45**, 153-159.

CHRISTENSEN A.H. & SCHUBERT K.R. (1983). Identification of a *Rhizobium trifolii* plasmid coding for nitrogen fixation and nodulation genes and its interaction with pJB5JI, a *Rhizobium leguminosarum* plasmid. Journal of Bacteriology. **156**, 592-599.

COOPER J.E., BJOURSON A.J. & THOMPSON J.K. (1987). Identification of *Lotus* rhizobia by direct DNA hybridisation of crushed root nodules. Applied and Environmental Microbiology. **53**, 1705-1707.

COURNOYER B., GOUY M. & NORMAND P. (1993). Molecular phylogeny of the symbiotic Actinomycetes of the genus *Frankia* matches host plant infection processes. Molecular Biology and Evolution. **10**, 1303-1316.

COURNOYER B. & NORMAND P. (1994). Characterization of a spontaneous thiostrepton resistant *Frankia alni* infective isolate using PCR-RFLP of *nif* and *gln*II genes. Soil Biology and Biochemistry. **26**, 553-559.

COUTINHO H.L.C., HANDLEY B.A., KAY H.E., STEVENSON L. & BERINGER J.E. (1993). The effect of colony age on PCR fingerprinting. Letters in Applied Microbiology. **17**, 282-284.

CRUICKSHANK R.H. & PITT J.I. (1987). Identification of species in *Penicillium* subgenus *penicillium* by enzyme electrophoresis. Mycologia. **79**, 614-620.

CUSHWA W.T. & MEDRANO J.F. (1996). Applications of the random amplified polymorphic DNA (RAPD) assay for genetic-analysis of livestock species. Animal Biotechnology. 7, 11-31.

DAIRE X., BOUDON-PADIEU E., BERVILLE A., SCHNEIDER B. & CAUDWELL A. (1992). Cloned DNA probes for detection of grapevine Flavescence dorée mycoplasma-like organism (MLO). Annals of Applied Biology. **121**, 95-103.

DE BRUIJN F.J. (1992). Use of repetitive (Repetitive Extragenic Palindromic Intergeneric Consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. Applied and Environmental Microbiology. **58**, 2180-2187.

DE JONG T.M., BREWIN N.J. JOHNSTON A.W.B. & PHILLIPS D.A. (1982). Improvements of symbiotic properties in *Rhizobium leguminosarum* by plasmid transfer. Journal of General Microbiology. **128**, 1829-1838.

DELAJUDIE P., WILLEMS A., POT B., DEWETTINCK D., MAESTROJUAN G., NEYRA M., COLLINS M.D., DREYFUS B., KERSTERS K. & GILLIS M. (1994). Polyphasic taxonomy of rhizobia - Emendation of the genus *Sinorhizobium* and description of *Sinorhizobium meliloti* comb nov, *Sinorhizobium saheli* sp nov, and *Sinorhizobium teranga* sp nov. International Journal of Systematic Bacteriology. **44**, 715-733.

DE LOS REYES-GAVILÁN C.G., LIMSOWTIN G.K.Y., TAILLÉZ P., SÉCHAUD L. & ACCOLAS J-P. (1992). A *Lactobacillus helveticus*-specific DNA probe detects restriction fragment length polymorphisms in this species. Applied and Environmental Microbiology. **58**, 3429-3432.

DEMEZAS D.H., REARDON T.B., WATSON J.M. & GIBSON A.H. (1991). Genetic diversity among *Rhizobium leguminosarum* biovar *trifolii* strains revealed by allozyme and restriction fragment length polymorphism analysis. Applied and Environmental Microbiology. **57**, 3489-3495.

DEMEZAS D.H., REARDON T.B., STRAIN S.R., WATSON J.M. & GIBSON A.H. (1995). Diversity and genetic structure of a natural population of *Rhizobium leguminosarum* bv *trifolii* isolated from *Trifolium subterraneum* L. Molecular Ecology. **4**, 209-220.

DENAIRE J., BOISTARD P., CASSE-DELBART F., ATHERLY A.G., BERRY J.O. & RUSSELL P. (1981). Indigenous plasmids of *Rhizobium*. International Review of Cytology. **13**, 225-246.

DIGBY P.G.N. & KEMPTON R.A. (1987). Multivariate analysis of ecological communities. Chapman and Hall, London and New York.

DJORDJEVIC M.A., ZURKOWSKI W. & ROLFE B.G. (1982). Plasmids and stability of symbiotic properties of *Rhizobium trifolii*. Journal of Bacteriology. **151**, 560-568.

DJORDJEVIC M.A., ZURKOWSKI W., SHINE J. & ROLFE B.G. (1983). Sym plasmid transfer to various symbiotic mutants of *Rhizobium trifolii*, *R. leguminosarum* and *R. meliloti*. Journal of Bacteriology. **156**, 1035-1045.

DOBEREINER J., DAY J.M., MASKALL S. & DART P. (1972). Rothamsted Experimental Station Report for 1971, part 1, page 97.

DOBROWOLSKI M.P. & O'BRIEN P.A. (1993). Use of RAPD-PCR to isolate a species specific DNA probe for *Phytophthora cinnamomi*. FEMS Microbiology Letters. **113**, 43-48.

DOOLEY J.J., HARRISON S.P., MYTTON L.R., DYE M., CRESSWELL A., SKØT L. & BEECHING J.R. (1993). Phylogenetic grouping and identification of *Rhizobium* isolates on the basis of random amplified polymorphic DNA profiles. Canadian Journal of Microbiology. **39**, 665-673.

DUGHRI M.H. & BOTTOMLEY P.J. (1983). Complementary methodologies to delineate the composition of *Rhizobium trifolii* populations in root nodules. Soil Science Society American Journal. **47**, 939-945.

DUGHRI M.H. & BOTTOMLEY P.J. (1984). Soil acidity and the composition of an indigenous population of *Rhizobium trifolii* in nodules of different cultivars of *Trifolium subterraneum* L. Soil Biology and Biochemistry. **16**, 405-411.

DYE M., SKØT L., MYTTON L.R., HARRISON S.P., DOOLEY J.J. & CRESSWELL A. (1995). A study of *Rhizobium leguminosarum* biovar *trifolii* populations from soil extracts using randomly amplified polymorphic DNA profiles. Canadian Journal of Microbiology. **41**, 336-344.

EARDLY B.D., MATERON L.A., SMITH N.H., JOHNSON D.A., RUMBAUGH M.D. & SELANDER R.K. (1990). Genetic structure of natural populations of the nitrogen-fixing bacterium *Rhizobium meliloti*. Applied and Environmental Microbiology. **56**, 187-194.

EARDLY B.D., YOUNG J.P.W. & SELANDER R.K. (1992). Phylogenetic position of *Rhizobium* sp. strain Or 191, a symbiont of both *Medicago sativa* and *Phaseolus vulgaris*, based on partial sequences of the 16S rRNA and *nif*H genes. Applied and Environmental Microbiology. **58**, 1809-1815.

ECKHARDT T. (1978). A rapid method for the identification of plasmid desoxyribonucleic acid in bacteria. Plasmid. **1**, 584-588.

ELKAN G.H. (1992). Taxonomy of the rhizobia. Canadian Journal of Microbiology. **38**, 446-450.

ENGVILD K.C., JENSEN E.S. & SKØT L. (1990). Parallel variation in isoenzyme and nitrogen fixation markers in a *Rhizobium* population. Plant and Soil. **128**, 283-286.

ENGVILD K.C. & NEILSEN G. (1985). Strain identification in *Rhizobium* by starch gel electrophoresis of isoenzymes. Plant and Soil. **87**, 251-256.

EXCOFFIER L., SMOUSE P.E. & QUATTRO J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes - Applications to human mitochondrial DNA restriction data. Genetics. **131**, 479-491.

EY P.L., KHANNA K., ANDREWS R.H., MANNING P.A. & MAYRHOFER G. (1992). Distinct genetic groups of *Giardia intestinalis* distinguished by restriction fragment length polymorphisms. Journal of General Microbiology. **138**, 2629-2637.

FEKETE A., BANTLE J.A., HALLING S.M. & STICH R.W. (1992). Amplification fragment length polymorphism in *Brucella* strains by use of polymerase chain reaction with arbitrary primers. Journal of Bacteriology. **174**, 7778-7783.

FOTTRELL P.F. & O'HORA A. (1969). Multiple forms of D(-)3-Hydroxybutyrate dehydrogenase in *Rhizobium*. Journal of General Microbiology. **57**, 287-292.

FREDRICKSON J.K., BEZDICEK D.F., BROCKMAN F.J. & LI S.W. (1988). Enumeration of Tn5 mutant bacteria in soil by using a most-probable-number-DNA hybridisation procedure and antibiotic resistance. Applied and Environmental Microbiology. **54**, 446-453.

FUJIMORI F. & OKUDA T. (1994). Application of the random amplified polymorphic DNA using the polymerase chain reaction for efficient elimination of duplicate strains in microbial screening .1. Fungi. Journal Of Antibiotics. **47**, 173-182.

FUQUAY J.I., BOTTOMLEY P.J. & JENKINS M.B. (1984). Complementary methods for the differentiation of *Rhizobium meliloti* isolates. Applied and Environmental Microbiology. **47**, 663-669.

GABRIEL D.W., HUNTER J.E., KINGSLEY M.T., MILLER J.W. & LAZO G.R. (1988). Clonal population structure of *Xanthomonas campestris* and genetic diversity among citrus canker strains. Molecular Plant-Microbe Interactions. **1,** 59-65.

GEMMELL G. & ROUGHLEY R.J. (1975). Report on the Australian Inoculants and Control Service. Proceedings of the 5th Australian legume nodulation conference, Brisbane. Paper 18.

GENIAUX E., LAGUERRE G. & AMARGER N. (1993). Comparison of geographically distant populations of *Rhizobium* isolated from root-nodules of *Phaseolus-vulgaris*. Molecular Ecology. **2**, 295-302.

GIBSON A.H. (1959). A study on the physiology and cytology of nodule formation on members of the family Leguminosae. Ph.D. Thesis, University of London.

GIBSON A.H. & NUTMAN P.S. (1960). Studies on the physiology of nodule formation. VII. A reappraisal of the effect of preplanting. Annals of Botany. **24**, 420-433.

GIBSON A.H., DEMEZAS D.H., GAULT R.R., BHUVANESWARI T.V. & BROCKWELL J. (1990). Genetic stability in Rhizobia in the field. Plant and Soil. **129**, 37-44.

GLYNN P., HIGGINS P., SQUARTINI A. & O'GARA F. (1985). Strain identification in *Rhizobium trifolii* using DNA restriction analysis and intrinsic antibiotic resistances. FEMS Microbiology Letters. **30**, 177-182.

GOODWIN P.H. & ANNIS S.L. (1991). Rapid identification of genetic variation and pathotype of *Leptosphaeria maculans* by random amplified polymorphic DNA assay. Applied and Environmental Microbiology. **57,** 2482-2486.

GOWER J.C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika. **53**, 325-338.

GUSS A. & DOBEREINER I. (1972). Effects of mineral nitrogen and soil temperature on nitrogen fixation of field beans (*Phaseolus vulgaris* L.). Pesq. Agropec. Bras., Ser. Agron. 7, 87-92.

HAIG S.M., RHYMER J.M. & HECKEL D.G. (1994). Population differentiation in randomly amplified polymorphic DNA of red-cockaded woodpeckers *Picoides borealis*. Molecular Ecology. **3**, 581-593.

HARDARSON G. & JONES D.G. (1979). The inheritance of preference for strains of *Rhizobium trifolii* by white clover (*Trifolium repens*). Annals of Applied Biology. **92**, 329-333.

HARRISON S.P., YOUNG J.P.W. & JONES D.G. (1987). *Rhizobium* population genetics: effect of clover variety and inoculum dilution on the genetic diversity sampled from natural populations. Plant and Soil. **103**, 147-150.

HARRISON S.P., JONES D.G., SCHUNMANN P.H.D., FORSTER J.W. & YOUNG J.P.W. (1988). Variation in *Rhizobium leguminosarum* bv *trifolii* Sym plasmids and the association with effectiveness of nitrogen fixation. Journal of General Microbiology. **134**, 2721-2730.

HARRISON S.P., JONES D.G. & YOUNG J.P.W. (1989a). *Rhizobium* population genetics: Genetic variation within and between populations from diverse locations. Journal of General Microbiology. **135**, 1061-1069.

HARRISON S.P., YOUNG J.P.W. & JONES D.G. (1989b). *Rhizobium* populations genetics: Host preference and strain competition effects on the range of *Rhizobium leguminosarum* bv *trifolii* genotypes isolated from natural populations. Soil Biology and Biochemistry. **21**, 981-986.

HARRISON S.P., MYTTON L.R., SKØT L., DYE M. & CRESSWELL A. (1992). Characterisation of *Rhizobium* isolates by amplification of DNA polymorphisms using random primers. The Canadian Journal of Microbiology. **38**, 1009-1015.

HARTMANN A. & AMARGER N. (1991). Genotypic diversity of an indigenous *Rhizobium meliloti* field population assessed by plasmid profiles, DNA fingerprinting, and insertion sequence typing. Canadian Journal of Microbiology. **37**, 600-608.

HARTMANN A., CATROUX G. & AMARGER N. (1992). *Bradyrhizobium japonicum* strain identification by RFLP analysis using the repeated sequence RS-Alpha. Letters in Applied Microbiology. **15**, 15-19.

HELZ G.E., BALDWIN I.L. & FRED E.B. (1927). Strain variation and host specificity of the root-nodule bacteria of the pea group. Journal of Agricultural Research (Washington). **35**. 1039-1055.

HIGGINS D.G. (1992). Sequence ordinations: a multivariate analysis approach to analysing large sequence data sets. CABIOS. **8**, 15-22.

HIRSCH P.R., JONES M.J., MCGRATH S.P. & GILLER K.E. (1993). Heavy-metals from past applications of sewage-sludge decrease the genetic diversity of *Rhizobium leguminosarum* biovar *trifolii* populations. Soil Biology and Biochemistry. **25**, 1485-1490.

HIRSCH P.R., VAN MONTAGU M., JOHNSTON A.W.B., BREWIN N.J. & SCHELL J. (1984). Physical identification of bacteriocinogenic, nodulation and other plasmids in strains of *Rhizobium leguminosarum*. Journal of General Microbiology. **120**, 403-412.

HODGSON A.L.M. & ROBERTS W.P. (1983). DNA colony hybridisation to identify *Rhizobium* strains. Journal of General Microbiology. **129**, 207-212.

HOLBEN W.E., JANSSON J.K., CHELM B.K. & TIEDJE J.M. (1988). DNA probe method for the detection of specific microorganisms in the soil bacterial community. Applied and Environmental Microbiology. **54**, 703-711.

HOPKINS B., O'CONNELL F.M. & HOPKINS J. (1991). Use of DNA fingerprinting in Paternity analysis of closely-related Exmoor ponies. Equine Veterinary Journal. **23**, 277-279.

HUGENHOLTZ P., STACKEBRANDT E. & FUERST J.A. (1993). A phylogenetic analysis of the genus *Blastobacter* with a view to its future reclassification. Systematic and Applied Microbiology. **17**, 51-57.

HUMPHREY B. & VINCENT J.M. (1969). The somatic antigens of two strains of *R. trifolii*. Journal of General Microbiology. **59**, 411-415.

IRELAN N.A & MEREDITH C.P. (1996). Genetic analysis of *Agrobacterium tumefaciens* and *A. vitis* using randomly amplified polymorphic DNA. American Journal of Enology and Viticulture. **47**, 145-151.

ISLAM R. & DART P.J. (1973). Effects of temperature and day length on the symbiosis of some tropical grain legumes. Rothamsted Experimental Station Report for 1972, part 1, page 85.

JACCARD P. (1901). Etude comparative de la distribution florale dans une portion des Alpes et des Jura. Bull. Soc. Vaudoise Sci. Nat. **37**, 547-579.

JAMANN S., FERNANDEZ M.P. & NORMAND P. (1993). Typing method for N2 fixing bacteria based on PCR-RFLP - Application to the characterization of *Frankia* strains. Molecular Ecology. **2**, 17-26.

JAMES F.C. & McCULLOCH C.E. (1990). Multivariate analysis in ecology and systematics: Panacea or Pandora's box?. Annual Review of Ecological Systems. **21**, 129-166.

JARVIS B.D.W., PANKHURST C.E. & PATEL J.J. (1982). *Rhizobium loti*, a new species of legume root nodule bacteria. International Journal of Systematic Bacteriology. **32**, 378-380.

JARVIS B.D.W., DOWNER H.L. & YOUNG J.P.W. (1992). Phylogeny of fast-growing soybean-nodulating rhizobia supports synonymy of *Sinorhizobium* and *Rhizobium* and assignment to *Rhizobium fredii*. International Journal of Systematic Bacteriology. **42**, 93-96.

JOHNSTON A.W.B. & BERINGER J.E. (1977). Chromosomal recombination between *Rhizobium* species. Nature (London). **267**, 611-613.

JONES D.G. & MORLEY S.J. (1981). The effect of pH on host plant 'preference' for strains of *Rhizobium trifolii* using fluorescent ELISA for strain identification. Annals of Applied Biology. **97**, 183-190.

JOUANIN L., DE LAJUDIE P., BAZETOUX S. & HUGUET T. (1981). DNA sequence homology in *Rhizobium meliloti* plasmids. Molecular and General Genetics. **182**, 189-195.

JUDD A.K., SCHNEIDER M., SADOWSKY M.J. & DE BRUIJN F.J. (1993). Use of repetitive sequences and the polymerase chain reaction technique to classify genetically related *Bradyrhizobium japonicum* serocluster 123 strains. Applied and Environmental Microbiology. **59**, 1702-1708.

KALNINS A. (1938). Fixation and mineralisation of atmospheric nitrogen. I. The efficiency of different strains of lupine and serradella nodule bacteria in fixing atmospheric. Acta Universitatis Latviensis. **49**.

KAMBHAMPATI S., BLACK W.C. & RAI K.S. (1992). Random amplified polymorphic DNA of mosquito species and populations (*Diptera, Culicidae*)-techniques, statistical-analysis, and applications. Journal Of Medical Entomology. **29**, 939-945.

KAMICKER B.J. & BRILL W.J. (1986). Identification of *Bradyrhizobium japonicum* nodule isolates from Wisconsin soybean farms. Applied and Environmental Microbiology. **51**, 487-492.

KISHINEVSKY B. & BAR-JOSEPH M. (1978). *Rhizobium* strain identification in *Arachis hypogeae* nodules by enzyme-linked immunosorbent assay (ELISA). Canadian Journal of Microbiology. **24**, 1537-1543.

KONDOROSI A., VINCZE E., JOHNSTON A.W.B. & BERINGER J.E. (1980). A comparison of three *Rhizobium* linkage maps. Molecular and General Genetics. **178**, 403-408.

KOSIER B., PUHLER A. & SIMON R. (1993). Monitoring the diversity of *Rhizobium meliloti* field and microcosm isolates with a novel rapid genotyping method using insertion elements. Molecular Ecology. **2**, 35-46.

KROL A.J.M., HONTELEZ J.G.J. & VANKAMMEN A. (1982). Only one of the large plasmids in *Rhizobium leguminosarum* strain *pre* is strongly expressed in the endosymbiotic state. Journal of General Microbiology. **128**, 1839-1847.

KUYKENDALL L.D., ROY M.A., O'NEILL J.J. & DEVINE T.E. (1988). Fatty acids, antibiotic resistance, and deoxyribonucleic acid homology groups of *Bradyrhizobium japonicum*. International Journal of Systematic Bacteriology. **38**, 358-361.

KUYKENDALL L.D., SWELIM D.M., HASHEM F.M., ABDELWAHAB S.M. & HEGAZI N.I. (1996). Symbiotic competence, genetic diversity and plasmid profiles of Egyptian isolates of a *Rhizobium* species from *Leucaena leucocephala* (lam) dewit. Letters in Applied Microbiology. **22**, 347-352.

LABES G., ULRICH A. & LENTZSCH P. (1996). Influence of bovine slurry deposition on the structure of nodulating *Rhizobium leguminosarum* bv *viciae* soil populations in a natural habitat. Applied and Environmental Microbiology. **62**, 1717-1722.

LAGUERRE G., MAZURIER S.I. & AMARGER N. (1992a). Plasmid profiles and restriction fragment length polymorphism of *Rhizobium leguminosarum* biovar *viceae* in field populations. FEMS Microbiology Ecology. **101**, 17-26.

LAGUERRE G., GENIAUX E., MAZURIER S.I., CASARTELLI R.R. & AMARGER N. (1992b). Conformity and diversity among field isolates of *Rhizobium leguminosarum* bv. *viciae*, bv. *trifolii*, and bv. *phaseoli* revealed by DNA hybridisation using chromosome and plasmid probes. Canadian Journal of Microbiology. **39**, 412-419.

LAGUERRE G., ALLARD M.R., REVOY F. & AMARGER N. (1994). Rapid identification of Rhizobia by restriction-fragment-length-polymorphism analysis of PCR-amplified 16S ribosomal-RNA genes. Applied and Environmental Microbiology. **60**, 56-63.

LAGUERRE G., MAVINGUI P., ALLARD M.R., CHARNAY M.P., LOUVRIER P., MAZURIER S.I., RIGOTTIERGOIS L. & AMARGER N. (1996). Typing of Rhizobia by PCR DNA-fingerprinting and PCR- restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions - application to *Rhizobium leguminosarum* and its different biovars. Applied and Environmental Microbiology. **62**, 2029-2036.

LAROCHE A., GAUDET D.A., SCHAALJE G.B., ERICKSON R.S. & GINNS T. (1995). Grouping and identification of low-temperature *Basidiomycetes* using mating, RAPD and RFLP analyses. Mycological Research. **99**, 297-310.

LECHEVALIER M.P. (1994). Taxonomy of the genus *Frankia* (Actinomycetales). International Journal of Systematic Bacteriology. **44**, 1-8.

LEE I-M., GUNDERSON D.E., DAVIS R.E. & CHIYKOWSKI L.N. (1992). Identification and analysis of a genomic strain cluster of Mycoplasmalike organisms associated with Canadian peach (eastern) X disease, western X disease, and clover yellow edge. Journal of Bacteriology. **174**, 6694-6698.

LINNE VON BERG K.-H. & BOTHE H. (1992). The distribution of denitrifying bacteria in soils monitored by DNA-probing. FEMS Microbiology Ecology. **86**, 331-340.

LONVAUD-FUNEL A., JOYEUX A. & LEDOUX O. (1991). Specific enumeration of lactic acid bacteria in fermenting grape must and wine by colony hybridisation with non-isotopic DNA probes. Journal of Applied Bacteriology. **71**, 501-508.

MADRZAK C.J., GOLINSKA B., KROLICZAK J., PUDELKO K., LAZEWSKA D., LAMPKA B. & SADOWSKY M.J. (1995). Diversity among field populations of *Bradyrhizobium japonicum* in Poland. Applied and Environmental Microbiology. **61**, 1194-1200.

MANULIS S. (1992). Evaluation of a DNA probe for the detection of *Erwinia herbicola* strains pathogenic on *Gypsophila paniculata*. Plant Pathology. **41**, 342-347.

MARTENSSON A.M., GUSTAFSSON J-G. & LJUNGGREN H.D. (1984). A modified, highly sensitive enzyme-linked immunosorbent assay (ELISA) for *Rhizobium meliloti* strain identification. Journal of General Microbiology. **130**, 247-253.

MARTENSSON A.M., GUSTAFSSON J-G. & LJUNGGREN H.D. (1987). Competition in soil between effective and ineffective strains of *Rhizobium leguminosarum* biovar *trifolii* in the nodulation of red clover, *Trifolium pratense* L., studied with ELISA. Journal of General Microbiology. **133**, 3467-3472.

MARTÍNEZ-ROMERO E. & ROSENBLUETH M. (1990). Increased bean (*Phaseolus vulgaris* L.) nodulation competitiveness of genetically modified *Rhizobium* strains. Applied and Environmental Microbiology. **56**, 2384-2388.

MARTÍNEZ-ROMERO E., SEGOVIA L., MERCANTE F.M., FRANCO A.A., GRAHAM P. & PARDO M.A. (1991). *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* spp. trees. International Journal of Systematic Bacteriology. **41**, 417-426.

MARTÍNEZ-ROMERO E. (1994). Recent developments in *Rhizobium* taxonomy. Plant and Soil. **161**, 11-20.

MAY S.N. & BOHLOOL B.B. (1983). Competition among *Rhizobium leguminosarum* strains for nodulation of lentils (*Lens esculenta*). Applied and Environmental Microbiology. **45**, 960-965.

Mc LAUGHLIN T.J., BORDELEAU L.M. & DUNICAN L.K. (1984). Competition studies with *Rhizobium trifolii* in a field experiment. Journal of Applied Bacteriology. **56**, 131-135.

MEADE J., HIGGINS P. & O'GARA F. (1985). Studies on the inoculation and competitiveness of a *Rhizobium leguminosarum* strain in soils containing indigenous Rhizobia. Applied and Environmental Microbiology. **49**, 899-903.

MEANS U.M, JOHNSON H.W. & DATE R.A. (1964). Quick serological method of classifying strains of *Rhizobium japonicumi* in nodules. Journal of Bacteriology. **3**, 547-553.

MEYER M.C. & PUEPPKE S.G. (1980). Differentiation of *Rhizobium japonicum* strain derivatives by antibiotic sensitivity patterns, lectin binding, and utilization of biochemicals. Canadian Journal of Microbiology. **26**, 606-612.

MIELENZ J.R., JACKSON L.E., O'GARA F. & SHANMUGAM K.T. (1979). Fingerprinting bacterial chromosomal DNA with restriction endonuclease *Eco*R1: comparison of *Rhizobium* spp. and identification of mutants. Canadian Journal of Microbiology. **25**, 803-807.

MINAMISAWA K., SEKI T., ONODERA S., KUBOTA M. & ASAMI T. (1992). Genetic relatedness of *Bradyrhizobium japonicum* field isolates as revealed by repeated sequences and various other characteristics. Applied and Environmental Microbiology. **58**, 2832-2839.

MINK M. & SIK T. (1983). Purification and electron-microscopy of a large plasmid of *Rhizobium meliloti* 41. Acta Microbiologica Hungarica. **30**, 79-82.

MOAWAD H.A., ELLIS W.R. & SCHMIDT E.L. (1984). Rhizosphere response as a factor in competition among three serogroups of indigenous *Rhizobium japonicum* for nodulation of field-grown soybeans. Applied and Environmental Microbiology. **47**, 607-612.

MORGAN U.M., CONSTANTINE C.C., ODONOGHUE P., MELONI B.P., OBRIEN P.A. & THOMPSON R.C.A. (1995). Molecular characterization of *Cryptosporidium* isolates from humans and other animals using random amplified polymorphic DNA analysis. American Journal Of Tropical Medicine And Hygiene. **52**, 559-564.

MORLEY S.J. & JONES D.G. (1980). A note on a highly sensitive modified ELISA technique for *Rhizobium* strain identification. Journal of Applied Bacteriology. **49**, 103-109.

MOZO T., CABRERA E. & RUIZ-ARGUESO T. (1988). Diversity of plasmid profiles and conservation of symbiotic nitrogen fixation genes in newly isolated *Rhizobium* strains nodulating Sulla (*Hedysarum coronarium* L.). Applied and Environmental Microbiology. **54**, 1262-1267.

MULLIS K., FALOONA F., SCHARF S., SAIKI R., HORN G. & ERLICH HA. (1986). Specific enzymatic amplification of DNA invitro - The polymerase chain reaction. Cold Spring Harbor Symposia on Quantitative Biology. **51**, 263-273.

MURPHY P.M. & MASTERSON C.L. (1970). Determination of multiple forms of esterases in *Rhizobium* by paper electrophoresis. Journal of General Microbiology. **61**, 121-129.

MUSSER J.M., HEWLETT E.L., PEPPLER M.S. & SELANDER R.K. (1986). Genetic diversity and relationships in populations of *Bordetella* spp. Journal of Bacteriology. **166**, 230-237.

MYTTON L.R. & LIVESEY C.J. (1983). Specific and general effectiveness of *Rhizobium trifolii* populations from different agricultural locations. Plant and Soil. **73**, 299-305.

MYTTON L.R., Mc ADAM N.J. & PORTLOCK P. (1977). Enzyme polymorphism as an aid to identification of *Rhizobium* strains. Soil Biology and Biochemistry. **10**, 79-80.

NEI M. & LI W-H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Science. **76**, 5269-5273.

NICK G. & LINDSTROM K. (1994). Use of repetitive sequences and the polymerase chain-reaction to fingerprint the genomic DNA of *Rhizobium galegae* strains and to identify the DNA obtained by sonicating the liquid cultures and root-nodules. Systematic and Applied Microbiology. **17**, 265-273.

NODARI R.O., TSAI S.M., GILBERTSON R.L. & GEPTS P. (1993a). Towards an integrated linkage map of common bean .2. Development of an RFLP-based linkage map. Theoretical and Applied Genetics. **85**, 513-520.

NODARI R.O., TSAI S.M., GUZMAN P., GILBERTSON R.L. & GEPTS P. (1993b). Toward an integrated linkage map of common bean .3. Mapping genetic-factors controlling host-bacteria interactions. Genetics. **134**, 341-350.

NORMAND P., ORSO S., COURNOYER B., JEANNIN P., CHAPELON C., DAWSON J., EVTUSHENKO L. & MISRA A.K. (1996). Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family *Frankiaceae*. International Journal of Systematic Bacteriology. **46**, 1-9.

NOUR S.M., CLEYETMAREL J.C., NORMAND P. & FERNANDEZ M.P. (1995). Genomic heterogeneity of strains nodulating chickpeas (*Cicer arietinum* L.) and description of *Rhizobium mediterraneum* Sp Nov. International Journal of Systematic Bacteriology. **45**, 640-648.

NOVIKOVA N.I., PAVLOVA E.A. VOROBJEV N.I. & LIMESHCHENKO E.V. (1994). Numerical taxonomy of *Rhizobium* strains from legumes of the temperate zone. International Journal of Systematic Bacteriology. **44**, 734-742.

NUTI M.P., LEDEBOER A.M., LEPIDI A.A. & SCHILPEROORT R.A. (1977). Large plasmids in different *Rhizobium* species. Journal of General Microbiology. **100**, 241-248.

NUTMAN P.S. (1946). Variation within strains of clover nodule bacteria in the size of nodule produced and in the effectivity of the symbiosis. Journal of Bacteriology. **51**, 411-432.

NUTMAN P.S. (1959). Some observations on root-hair infection by nodule bacteria. Journal of Experimental Botany. **10**, 250-263.

OCHMAN H., WHITTAM T.S., CAUGANT D.A. & SELANDER R.K. (1983). Enzyme polymorphism and genetic population structure in *Escherichia coli* and *Shigella*. Journal of General Microbiology. **129**, 2715-2726.

OCHMAN H. & SELANDER R.K. (1984). Standard reference strains of *Escherichia coli* from natural populations. Journal of Bacteriology, **157**, 690-693.

O'DONOUGHUE L.S., SOUZA E., TANKSLEY S.D. & SORRELLS M.E. (1994). Relationships among North-American oat cultivars based on restriction-fragment-length-polymorphisms. Crop Science. **34**, 1251-1258.

OKUDA T., YANAGISAWA A., FUJIMORI F., NISHIZUKA Y., TAKEHANA Y. & SUGIYAMA M. (1995). New isolation methods and polymerase chain-reaction strain discrimination techniques for natural-products screening programs. Canadian Journal Of Botany-Revue Canadienne De Botanique. **73**, 946-954.

OLIVER M.A. & WEBSTER R. (1989). Statistical Methods in Soil and Land Resource Survey. Oxford University Press.

OXELMAN B. (1996). RAPD patterns, nrDNA its sequences and morphological patterns in silene section Sedoideae (*Caryophyllaceae*). Plant Systematics And Evolution. pp.93-116.

PAFFETTI D., SCOTTI C., GNOCCHI S., FANCELLI S. & BAZZICALUPO M. (1996). Genetic diversity of an Italian *Rhizobium meliloti* population from different *Medicago sativa* varieties. Applied And Environmental Microbiology. **62**, 2279-285.

PEAKALL R., SMOUSE P.E. & HUFF D.R. (1995). Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious Buffalograss *Buchloe dactyloides*. Molecular Ecology. **4**, 135-147.

PESSANHA G.G., FRANCO A.A., DOBEREINER J., GROSZMANN A. & BRITTO O.P.P.de S. (1972). Negative correlation of nodulation and grain production of beans (*Phaseolus vulgaris*) in soils where nitrogen is not a limiting factor of plant growth. Pesq. Agropec. Bras., Ser. Agron. 7, 49-56.

PICARD C., PONSONNET C., PAGET E., NESME X. & SIMONET P. (1992). Detection and enumeration of bacteria in soil by direct DNA extraction and polymerase chain reaction. Applied and Environmental Microbiology. **58**, 2717-2722.

PINERO D., MARTINEZ E. & SELANDER R.K. (1988). Genetic diversity and relationships among isolates of *Rhizobium leguminosarum* biovar *phaseoli*. Applied and Environmental Microbiology. **54**, 2825-2832.

PRAKASH R.K. & ATHERLY A.G. (1986). Plasmids of *Rhizobium* and their role in symbiotic nitrogen fixation. International Review of Cytology. **104**, 1-24.

ROBERT F.M. & SCHMIDT E.L. (1983). Population changes and persistence of *Rhizobium phaseoli* in soil and rhizospheres. Applied and Environmental Microbiology. **45**, 550-556.

SADOWSKY M.J., BOHLOOL B.B. & KEYSER H.H. (1987a). Serological relatedness of *Rhizobium fredii* to other Rhizobia and to the Bradyrhizobia. Applied and Environmental Microbiology. **53**, 1785-1789.

SADOWSKY M.J., TULLY R.E., CREGAN P.B. & KEYSER H.H. (1987b). Genetic diversity in *Bradyrhizobium japonicum* serogroup 123 and its relation to genotype-specific nodulation of soybean. Applied and Environmental Microbiology. **53**, 2624-2630.

SAWADA H., IEKI H., OYAIZU H. & MATSUMOTO S. (1993). Proposal for rejection of *Agrobacterium tumefaciens* and revised descriptions for the genus *Agrobacterium* and for *Agrobacterium radiobacter* and *Agrobacterium rhizogenes*. International Journal of Systematic Bacteriology. **43**, 694-702.

SCHAAD N.W. (1979). Serological identification of plant pathogenic bacteria. Annual Review of Phytopathology. **17**, 123-147.

SCHNEIDER M. & DE BRUIJN F.J. (1996). REP-PCR mediated genomic fingerprinting of Rhizobia and computer-assisted phylogenetic pattern-analysis. World Journal of Microbiology and Biotechnology. **12**, 163-174.

SCHOFIELD P.R., GIBSON A.H., DUDMAN W.F. & WATSON J.M. (1987). Evidence for genetic exchange and recombination of *Rhizobium* symbiotic plasmids in a soil population. Applied and Environmental Microbiology. **53**, 2942-2947.

SCHULZ T.F., BAUER C., LORENZ D., PLAPP R., EICHHORN K.W. (1993). Studies on the evolution of *Agrobacterium vitis* as based on genomic fingerprinting and IS element analysis. Systematic and Applied Microbiology. **16**, 322-329.

SEGOVIA L., PINERO D., PALACIOS R. & MARTÍNEZ-ROMERO E. (1991). Genetic structure of a soil population of nonsymbiotic *Rhizobium leguminosarum*. Applied and Environmental Microbiology. **57**, 426-433.

SELANDER R.K., CAUGANT D.A., OCHMAN H., MUSSER J.M., GILMOUR M.N. & WHITTAM T.S. (1986). Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Applied and Environmental Microbiology. **51**, 873-884.

SELANDER R.K., MUSSER J.M., CAUGANT D.A., GILMOUR M.N. & WHITTAM T.S. (1987). Population genetics of pathogenic bacteria. Microbial Pathogenesis. **3**, 1-7.

SELENSKAPOBELL S., GIGOVA L. & PETROVA N. (1995). Strain-specific fingerprints of *Rhizobium galegae* generated by PCR with arbitrary and repetitive primers. Journal of Applied Bacteriology. **79**, 425-431.

SELLSTEDT A., WULLINGS B., NYSTROM U. & GUSTAFSSON P. (1992). Identification of *Casuarina-Frankia* strains by use of polymerase chain reaction (PCR) with arbitrary primers. FEMS Microbiology Letters. **93**, 1-6.

SHI Y.L., LOOMIS P., CHRISTIAN D., CARRIS L.M. & LEUNG H. (1996). Analysis of the genetic-relationships among the wheat bunt fungi using RAPD and ribosomal DNA markers. Phytopathology. **86**, 311-318.

SIMON R., HOTTE B., KLAUKE B. & KOSIER B. (1991). Isolation and characterization of insertion sequence elements from gram-negative bacteria by using new broad-host-range, positive selection vectors. Journal of Bacteriology. **173**, 1502-1508.

SINCLAIR M.J. & EAGLESHAM A.R.J. (1984). Intrinsic antibiotic resistance in relation to colony morphology in three populations of west african cowpea Rhizobia. Soil Biology and Biochemistry. **16**, 247-251.

SOBERÓN-CHÁVEZ G., NÁJERA R., OLIVERA H. & SEGOVIA L. (1986). Genetic rearrangements of a *Rhizobium phaseoli* symbiotic plasmid. Journal of Bacteriology. **167**, 487-491.

SOBRAL B.W.S., HONEYCUTT R.J. & ATHERLY A.G. (1991). The genomes of the family *Rhizobiaceae*: Size, stability and rarely cutting restriction endonucleases. Journal of Bacteriology. **173**, 704-709.

SOUTHERN E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology. **98**, 503-517.

SOVINSKI S.M., BAIRD L.S., BUDOWLE B., CARUSO J.F., CHEEMA D.P.S., DUNCAN G.T., HAMBY P.P., MASIBAY A.S., SHARMA V.K. & TAHIR A. (1996). Development of a deoxyribonucleic-acid (DNA) restriction fragment length polymorphism (RFLP) database for Punjabis in East Punjab, India. Forensic Science International. **79**, 187-198.

SPITZBARTH M., PUHLER A. & HEUMANN W. (1979). Characterisation of plasmids isolated from *Rhizobium meliloti*. Archives of Microbiology. **121**, 1-7.

STANIER R.Y., INGRAHAM J.L., WHEELIS M.L. & PAINTER P.R. (1987). General microbiology, 5th edition. MacMillan Education Ltd, London.

STANLEY J., BROWN G.G. & VERMA D.P.S. (1985). Slow-growing *Rhizobium japonicum* comprises two highly divergent symbiotic types. Journal of Bacteriology. **163**, 148-154.

STEIN M., BROMFIELD E.S.P. & DYE M. (1982). An assessment of a method based on intrinsic antibiotic resistance for identifying *Rhizobium* strains. Annals of Applied Biology. **101**, 261-267.

STEWART C.N., ROSSON G., SHIRLEY B.W. & PORTER D.M. (1996). Population genetic-variation in rare and endangered Iliamna (*Malvaceae*) in Virginia. Biological Journal Of The Linnean Society. **58**, 357-369.

STEWART C.N. & EXCOFFIER L. (1996). Assessing population genetic-structure and variability with RAPD data - application to *Vaccinium macrocarpon* (American cranberry). Journal Of Evolutionary Biology. **9**, 153-171.

STREIT W., BJOURSON A.J., COOPER J.E. & WERNER D. (1993). Application of subtraction-hybridisation for the development of a *Rhizobium leguminosarum* biovar *phaseoli* and *Rhizobium tropici* group-specific DNA probe. FEMS Microbiology Ecology. **13**, 59-67.

SZALANSKI A.L., TAYLOR D.B. & PETERSON R.D. (1996). Population-genetics and gene variation of Stable fly populations (*Diptera, Muscidae*) in Nebraska. Journal of Medical Entomology. **33**, 413-420.

TANAKA H., SAWAIRI S. & OKUDA T. (1994). Application of the random amplified polymorphic DNA using the polymerase chain-reaction for efficient elimination of duplicate strains in microbial screening .3. Bacteria. Journal of Antibiotics. **47**, 194-200.

THEIN S.L. & WALLACE R.B. (1986). The use of synthetic oligonucleotides as specific hybridisation probes in the diagnosis of genetic disorders. Human Genetic Diseases: A Practical Approach. Edited by K.D. Davis. IRL Press, Herndon, VA. pp. 33-50.

THORNTON H.G. (1931). The cultivation and inoculation of lucerne. Journal of the Farmers Club, Part 2, pages 19-33.

TICHY H.V. & LOTZ W. (1981). Identification and characterisation of large plasmids in newly isolated strains of *Rhizobium leguminosarum*. FEMS Microbiology Letters. **10**, 203-207.

TSAI Y-L. & OLSON B.H. (1992). Detection of low numbers of bacterial cells in soils and sediments by polymerase chain reaction. Applied and Environmental Microbiology. **58**, 754-757.

TURCO R.F. & BEZDICEK D.F. (1987). Diversity within two serogroups of *Rhizobium leguminosrum* native to soils in the Palouse of Eastern Washington. Annals of Applied Biology. **111**, 103-114.

VANBERKUM P., KOTOB S.I., BASIT H.A., SALEM S., GEWAILY E.M. & ANGLE J.S. (1993). Genotypic diversity among strains of *Bradyrhizobium japonicum* belonging to serogroup-110. Applied and Environmental Microbiology. **59**, 3130-3133.

VANCOPPENOLLE B., McCOUCH S.R., WATANABE I., HUANG N. & VANHOVE C. (1995). Genetic diversity and phylogeny analysis of *Anabaena azollae* based on RFLPs detected in *Azolla anabaena azollae* DNA complexes using *nif* gene probes. Theoretical And Applied Genetics. **91**, 589- 597.

VANEECHOUTTE M., ROSSAU R., DE VOS P., GILLIS M., JANSSENS D., PAEPE N., DE ROUCK A., FIERS T., CLAEYS G. & KERSTERS K. (1992). Rapid identification of bacteria of the Comamonadaceae with amplified ribosomal DNA-restriction analysis (ARDRA). FEMS Microbiol. Letters. **93**, 227-234.

VAUTERIN L. & VAUTERIN P. (1992). Computer-aided objective comparison of electrophoresis patterns for grouping and identification of microorganisms. European Microbiology. **Nov/Dec**, 37-41.

WATSON J.M. & SCHOFIELD P.R. (1985). Species-specific, symbiotic plasmid located repeated DNA sequences in *Rhizobium trifolii*. Molecular and General Genetics. **199**, 279-289.

WELSH J. & McCLELLAND M. (1990). Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research. **18**, 7213-7218.

WELSH J. & McCLELLAND M. (1991). Genomic fingerprinting using arbitrarily primed PCR and a matrix of pairwise combinations of primers. Nucleic Acids Research. **19**, 5275-5279.

WHEATCROFT R. & WATSON R.J. (1987). Identification and characterization of insertion-sequence ISRm1 in rhizobium-meliloti JJ1C10. Canadian Journal of Microbiology. **33**, 314-321.

WHEATCROFT R. & WATSON R.J. (1988a). A positive strain identification method for *Rhizobium meliloti*. Applied and Environmental Microbiology. **54**, 574-576.

WHEATCROFT R. & WATSON R.J. (1988b). Distribution of insertion sequence ISRm1 in *Rhizobium meliloti* and other gram-negative bacteria. Journal of General Microbiology. **134**, 113-121.

WHITTAM T.S., OCHMAN H. & SELANDER R.K. (1983). Multilocus genetic structure in natural populations of *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences. **80**, 1751-1755.

WILKERSON R.C., PARSONS T.J., ALBRIGHT D.G., KLEIN T.A. & BRAUN M.J. (1993). Random amplified polymorphic DNA (RAPD) markers readily distinguish cryptic mosquito species (Diptera: Culicidae: *Anopheles*). Insect Molecular Biology. **1**, 205-211.

WILLEMS A. & COLLINS M.D. (1993). Phylogenetic analysis of Rhizobia and Agrobacteria based on 16S ribosomal RNA gene sequences. International Journal of Systematic Bacteriology. **43**, 305-313.

WILLIAMS J.G.K., KUBELIK A.R., LIVAK K.J., RAFALSKI J.A. & TINGEY S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research. **18,** 6531-6535.

WITTY J.F. & DAY J.M. (1977). Field inoculation of sweet lupins. Rothamsted Experimental Station Report for 1976, part 1, page 288.

WRIGHT S.F., FOSTER J.G. & BENNETT O.L. (1986). Production and use of monoclonal antibodies for identification of strains of *Rhizobium trifolii*. Applied and Environmental Microbiology. **52**, 119-123.

XU L.M., GE C., CUI Z., LI J. & FAN H. (1995). *Bradyrhizobium liaoningense* Sp. Nov., isolated from the root nodules of soybeans. International Journal of Systematic Bacteriology. **45**, 706-711.

YOUNG J.P.W. (1985). *Rhizobium* population genetics: Enzyme polymorphism in isolates from pea, clover, beans and lucerne grown at the same site. Journal of General Microbiology. **131**, 2399-2408.

YOUNG J.P.W., DEMETRIOU L. & APTE R.G. (1987). *Rhizobium* population genetics: Enzyme polymorphism in *Rhizobium leguminosarum* from plants and soil in a pea crop. Applied and Environmental Microbiology. **53**, 397-402.

YOUNG J.P.W. & WEXLER M. (1988). Sym plasmid and chromosomal genotypes are correlated in field populations of *Rhizobium leguminosarum*. Journal of General Microbiology. **134**, 2731-2739.

YOUNG J.P.W., DOWNER H.L. & EARDLY B.D. (1991). Phylogeny of the phototrophic *Rhizobium* strain BT Ail by polymerase chain reaction-based sequencing of a 16S rRNA gene segment. Journal of Bacteriology. **173**, 2271-2277.

YOUNG C.C., BURGHOFF R.L., KEIM L.G., MINAK-BERNERO V., LUTE J.R. & HINTON S.M. (1993). Polyvinylpyrrolidone-agarose gel electrophoresis purification of polymerase chain reaction-amplifiable DNA from soils. Applied and Environmental Microbiology. **59**, 1972-1974.

Appendices

Appendix 1

GENSTAT data analysis program:

```
open'DATAFILE.txt';channel=2 - DATAFILE.txt contains 1 and 0 data
units [nvalues=N] - N = the number of samples
text track
read [channel=2;prin=*]track,band[1...BN] - BN = number of bands
symmetricmatrix [rows=track]tracksim
fsimilarity [similarity=tracksim;print=similarities] band[1...BN];test=Simple
delete [redefine=y]band[1...BN]
hcluster [print=amalgamations,dendrogram;method=average] tracksim
lrv [rows=track;columns=4]latent;trackscr
pco [print=roots] tracksim;latent
variate score[1...4]
equate [oldformat=!((1,-3)N,-1)]trackscr;score
print track,score[1...4]
graph score[2];score[1];symbols=track
graph score[3];score[1];symbols=track
graph score[3];score[2];symbols=track
text labels;
!t('T','T','T','T','T','T','T','T','T','T','T','T','T','T','T','T','T','T','V','V','V','V','V','V','V','V','V','V','
V','V','V','V','V','V','V','V','V','V','P','P','P','P','P','P','P','P','P','P','P','P','P','P','P','P','P','P','M','M
','M','M','M','M','M','M','M','M','M','M','M','M','M','M','M','B','B','B','B','B','B','B','B','B','B'
,'B','B','B')
pen 1; method=point; symbols=labels; size=.5
dgraph [title='PCO plot ordinate1 v ordinate 2'] score[2];score[1]
open 'PLOTFILE.pcx' ; channel=3; filetype=graphics - plot saved as PLOTFILE.pcx
device 3
pen 1; method=point; symbols=labels; size=.5
dgraph [title='PCO plot ordinate1 v ordinate 2'] score[2];score[1]
close 3; filetype=graphics
device 1
axes 1; style=box
axes 1; yorigin=0; xorigin=0
pen 32; linestyle=1; thickness=0.75
axes 1; penaxes=32
dgraph [title='PLOT TITLE'] score[2];score[1]
```

Appendix 2

SPH1 band presence (1) and absence (0) data:

V1 0 0 0 0 0 0 1 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 1 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0
V2 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 1 0 0 1 0 1 0 0 0 0 0 0
V3 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0
V4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
V5 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
V6 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 1 1 0 0 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
V7 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 1 0 0 1 0 0 0 0
V8 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0
V9 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1
VA 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 1
VB 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0
VC 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0
VD 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 0 0 0 1 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
VE 0 0 0 1 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
VF 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
VG 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
VH 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 1 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
VI 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
T1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 1 0 1 0 1 0 0
T2 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0
T3 0 0 0 0 0 1 0 1 0 1 0 0 1 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0
T4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
T5 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 1 0 1 0 0 0 0
T6 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0
T7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0
T8 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0
T9 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0
TA 1 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 0
TB 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0
TC 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0
TD 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
TE 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0
TF 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0
TG 0 0 0 0 0 1 0 1 0 1 0 0 1 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
TH 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0

TI 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

P1 0 0 0 1 0 0 1 1 0 0 0 0 1 1 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0

P2 0 1 1 1 0 0 0 1 0 0 0 0 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

P3 0 0 1 1 0 0 0 1 0 1 0 0 1 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0

P4 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0

P5 0 0 0 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0

P6 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0

P7 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0

P8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0

P9 0 0 0 0 0 1 0 0 1 0 0 1 0 0 0 1 1 0 0 0 0 1 0 0 1 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0

PA 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0

PB 0 0 0 0 0 1 0 0 1 0 0 1 0 0 0 1 1 0 0 0 0 1 0 0 1 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0

PC 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0

PD 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 1 0 1 1 0 0 0 0 0 1 1 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0

PE 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0

PF 0 1 0 1 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0

PG 0 0 1 0 0 0 1 1 0 0 0 1 1 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0

PH 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 1 0

PI 0 0 1 0 0 0 1 0 1 0 0 1 1 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0

M1 0 0 0 1 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0

M2 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0

M3 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 1 0 0 0

M4 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0

M5 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 1 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0

M6 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0

M7 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0

M8 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0

M9 0 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0

MA 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0

MB 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0

MC 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0

MD 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 1 0 1 0 0 0

ME 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0

MF 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0

MG 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0

MH 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0

B1 0 0 1 0 1 0 0 0 0 0 0 0 0 1 1 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 1

B2 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

B3 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0

B4 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0

B5 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 1 0 0 1 1 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 1

B6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

B7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 0 1 0 1 0 0 0 0 0 1

B8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0

B9 0 0 0 0 0 1 0 1 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 0

BA 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 1 1 1 0 1 0 0 0 1 0 1

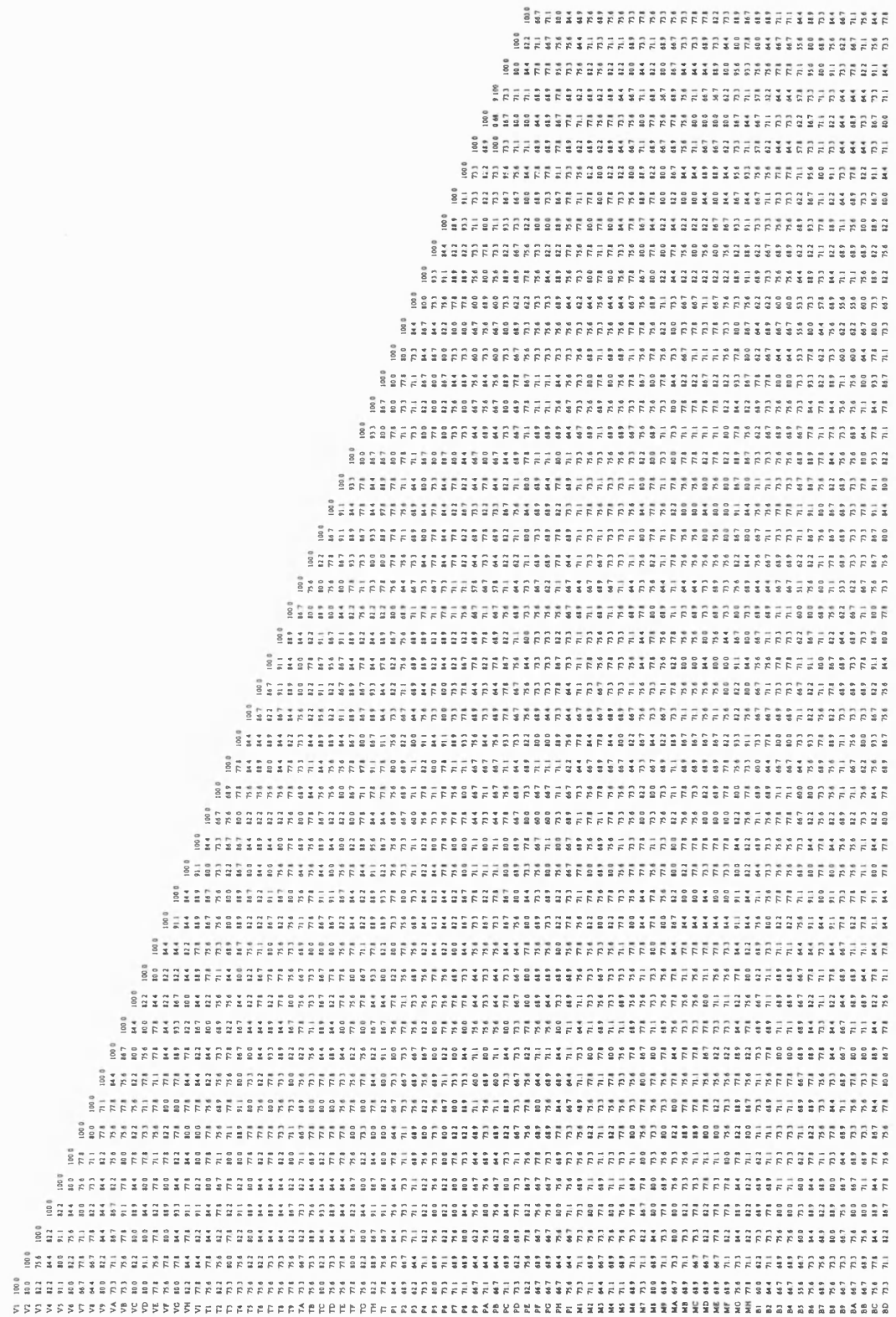
BB 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 0 0 1 0 0 0 1 1 0 0 0 0 0 0 0 0

BC 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

BD 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0:

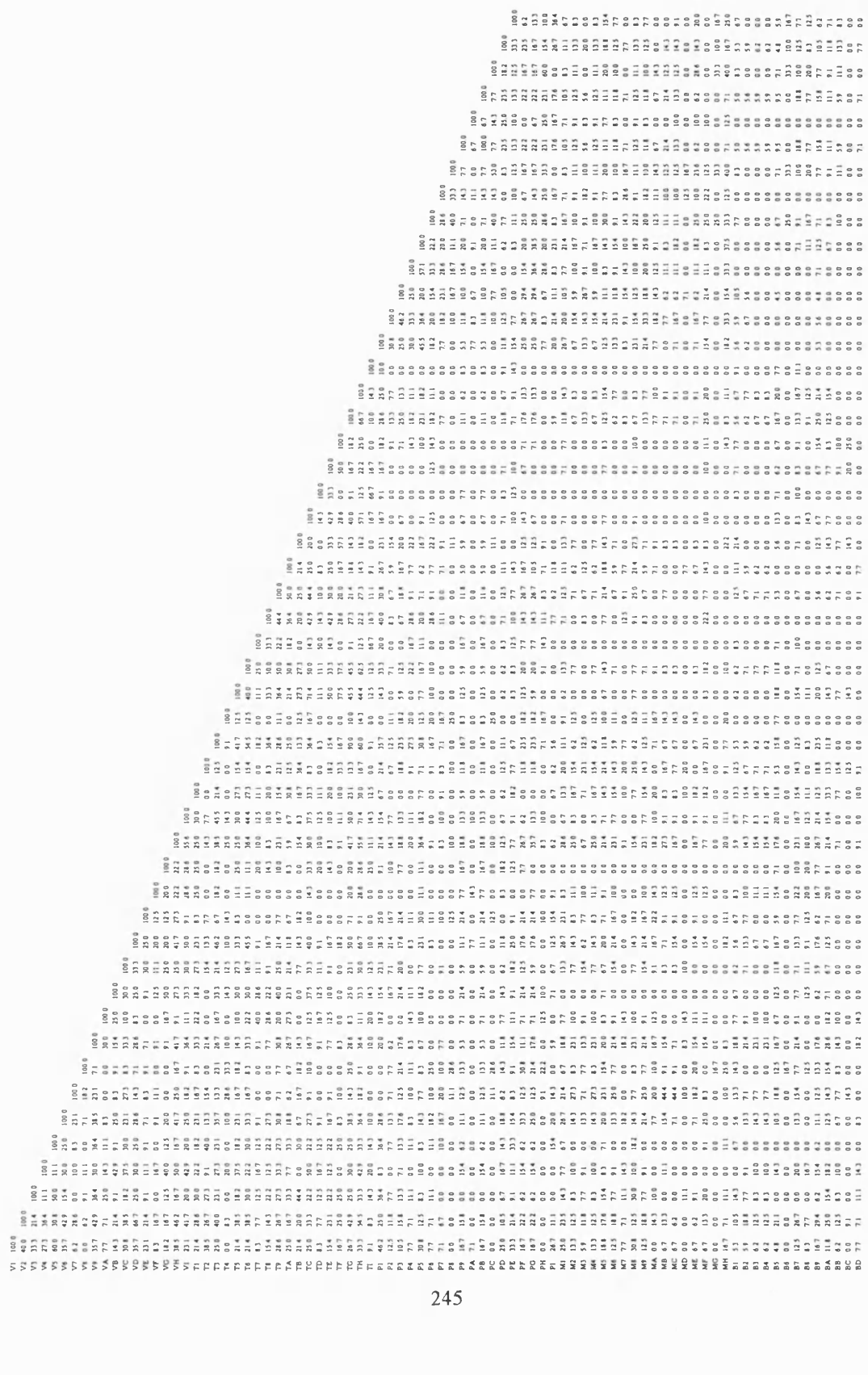
Appendix 3

Similarity matrix from SPH1 data produced using the simple matching coefficient:



Appendix 4

Similarity matrix from SPH1 data produced using the Jaccard matching coefficient:



Appendix 5

SPH3 presence (1) and absence (0) data:

T1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0

T2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0

T3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0

T4 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 1 1 0 0 0 0 0 1 0 0 0 0 0 0 1 0 1 0 1 1 0 1 0 1 0 0 0 0 1 1 0 1 0 0 0 0 0 0

T5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 1 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0

T6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0

T7 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 1 0 1 1 0 0 0 1 0 1 0 1 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1

T8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 1 0 0 1 0 1 0 0 0 1 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0

T9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

TA 0 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 1 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0

TB 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0 1 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0

TC 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 1 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0

TD 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 1 0 0 1 0 0 0 1 0 1 1 0 1 0 0 1 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0

TE 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 1 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

TF 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 1 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0

TG 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 1 1 0 1 1 0 1 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1

TH 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0 1 0 1 0 0 1 0 0 0 1 0 1 0 0 1 1 1 0 1 0 0 0 1 0 1 1 0 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0

TI 0000000000000000000000010010011010101000010100000101000001000
V1 0000000000000000000000010001000001010010100011001010000010000
V2 0000000000100000100001000001001100100000100010001001000000000
V3 0000000000000000100000100010001001100101001000010100000010000
V4 0000000000000010000000100100100010000101000010000000010000000
V5 0000000000000000000000000000100000000000000000000000000000000
V6 0000000000100001001000100000000010100101001100010100000010000
V7 0000000000000010001000100100100011110010100011000101000000000
V8 0000000000000000010000100100000111100101010100100101010000000
V9 0000000000000000100001001001000001010001001001000100100000000
VA 0000000000000000100001001001000001010001011001000000100000000
VB 0000000000000001000100001000000010010010001001000010100001000
VC 0000000001001000100010101010000100010010011010000010000001000
VD 0000000000100000100001000101000010010011000010000000000010000
VE 0000000000100000100001010011000010100000001000000010000100000
VF 0000000010100000000000010001000010100010111010000101010000000
VG 0000000000000000100100001001000101010001011101001000100000000
VH 0000000001010000000000001001000110110011101010000100000010000
VI 0000000000001000000000000100000100100101011000100001001000010000

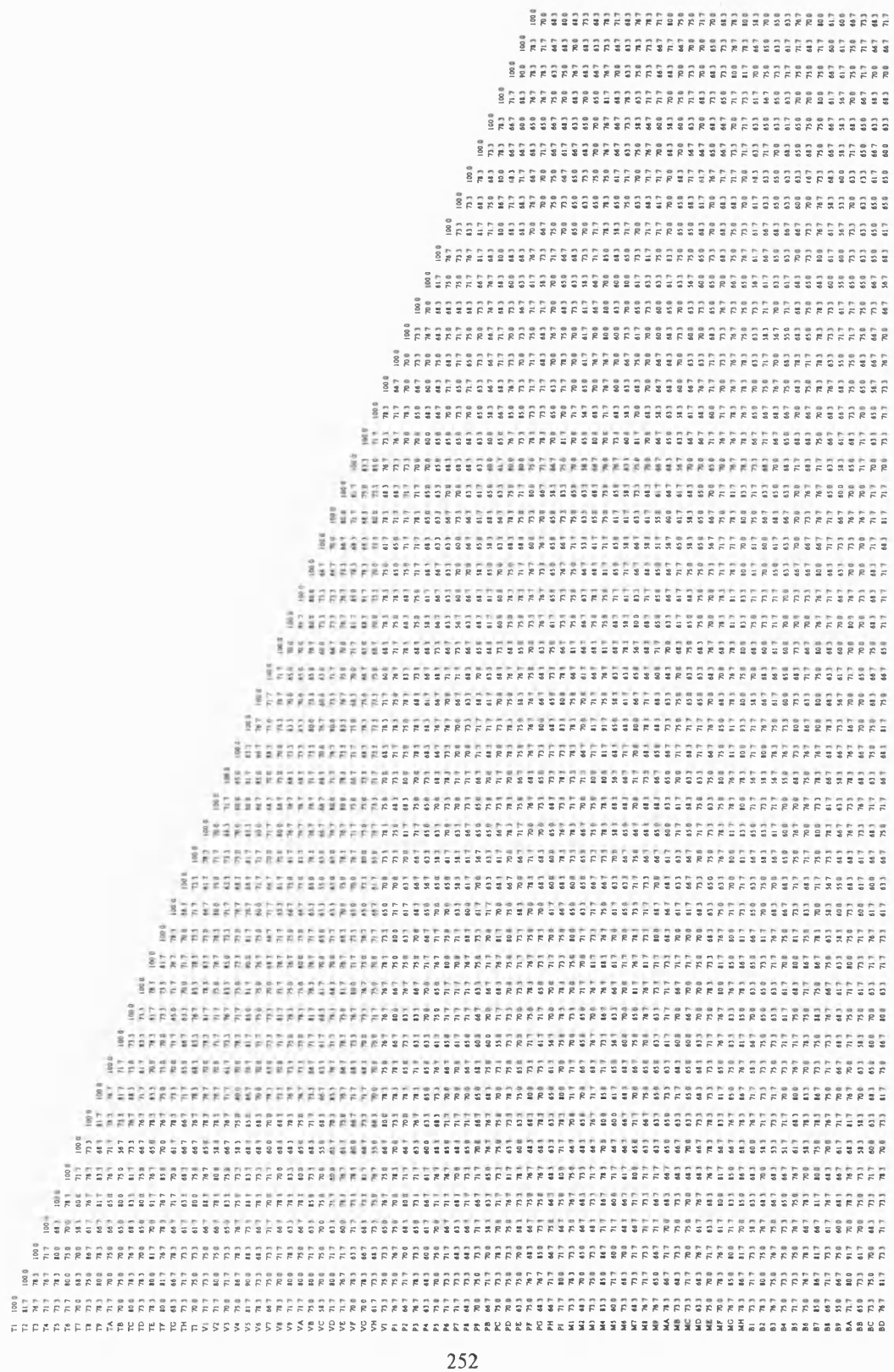
P1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 1 1 0 0 0 0 1 0 1 0 0 1 0 0 0 1 0 1 0 0 0
0 0 0 0 0 0 0 0
P2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 1 1 1 0 0 1 1 0 1 0 1 1 0 1 0 0 1 0 1 0 0 0
0 0 0 0 0 0 0 0
P3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 1 0 0 1 0 1 1 0 0 0 0 0 1 0 1
0 0 0 0 0 0 0 0
P4 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 1 0 0 0
1 0 1 0 0 0 0 0
P5 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 1 0 1 0 0 0 0 0 1 0 0 1 0 0 1 0 1 0 0 0 1 0 0 1 0 1 0 0 0 1 0 1 0 0 0 0
0 0 0 0 0 0 0 0
P6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0
1 0 0 1 0 0 0 0
P7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 1 0 1 0 0 1 1 0 1 0 1 0 1 1 1 1 0 1 0 0 0 1 0 0 1 0 0 0
1 0 0 0 0 0 0 0
P8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 0 1 1 0 1 0 0 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 0 0 1 0
0 0 0 0 0 0 0 0
P9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 1 0 1 0 1 0 0 1 1 0 0 0 0 1 0 0 0 1 1 0 1 0 1 1 0 0 1 0 1 0 0 0 0
0 1 0 0 0 0 0 0
PB 0 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 1 0 0 0 1 0 1 1 1 0 1 0 0 1 0 1 0 0 0 0
0 1 0 0 0 0 0 0
PC 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 0 1 0 0 1 1 0 1 0 1 0 1 0 0 1 0 1 1 0 0 1 0 1 0 0 0 0
0 0 0 0 0 0 0 0
PD 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 0 0 0 1 0 1 0 1 1 1 1 0 0 0 0 0 0 0
0 0 0 1 0 0 0 0
PE 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 1 0 1 0 0 0 0 0 0 0 0
0 0 0 1 0 0 0 0
PF 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1 0 1 0 1 0 1 0 1 0 0 0 1 1 0 1 0 1 1 0 0 1 0 1 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
PG 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 1 0 0 1 0 1 1 1 1 1 1 0 0 1 0 0 0 0
0 0 0 0 0 0 0 0
PH 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 1 0 1 0 0 1 1 0 1 0 0 0 1 1 1 1 1 1 0 0 1 0 1 0 0 1 1 0
0 0 0 0 0 0 0 0
PI 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 0
0 0 0 0 0 0 0 0
M1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 1 1 0 1 0 1 0 0 0 0 0 0 1 1 0
0 0 0 1 0 0 0 0
M2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 0 0 0 1 1 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1 1 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0
0 1 0 1 0 0 0 0

M3 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 1 0 1 0 0 1 1 0 0 0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 1 0
0 0 0 0 0 0 0 0
M4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
M5 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 1 0 1 1 1 0 1 0 0 0 1 0 1 1 1 1 0 0 1 0 0 0 1 1 0
0 1 0 0 0 0 1 0
M6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 1 1 0 1 0 0 1 1 0 1 1 1 0 1 0 1 0 0 0 0 1 1 0
1 0 1 0 0 0 0 0
M7 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 1 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
M8 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0 1 1 0 0 1 0 1 0 0 1 1 0 0 0 0 0 1
0 0 0 0 0 0 0 0
M9 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 0 0 1 0 1 0 0 0 1 0 1 1 0 0 1 0 1 1 0 1 0 0 1 0 0 1 0
1 0 1 0 0 0 0 0
MA 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 1 0 1 0 0 1 0 0 1 0 1 0 0 1 1 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0
1 0 0 0 0 0 0 0
MB 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1 1 0 0 0 0 1 1 0 1 0 1 0 0 1 0 0
0 0 1 0 0 0 0 0
MC 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 1 1 0 1 0 0 0 0 1 1 0
1 1 0 0 0 0 0 0
MD 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 1 0 1 0 0 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1 0 1 0 0
0 0 0 0 0 0 0 0
ME 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0 1 0 1 1 0 0 0 1 0 1 1 0 0 0 1 0 1 0 0 1 1 0 0 0 0 0 0 0 1 0
0 0 1 0 0 0 0 0
MF 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 1 1 0 1 0 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0
0 0 0 0 0 0 0 0
MG 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0
MH 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0
B1 0 0 0 0 0 1 0 0 0 1 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 1 1 0 0 0 1 1 0 1 0 0 1 0 0 1 0 1 1 0 0 0 1 0 0 1 0
1 0 0 0 0 0 0 0
B2 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0
0 0 1 0 0 0 0 0
B3 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0
1 0 1 0 0 0 0 0
B4 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 0 1 0 1 1 0
1 0 1 0 0 0 0 0

B5 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 1 0 1 0 0 0 0 1 0 0
0 0 0 0 0 0 0 0
B6 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0
B7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0
B8 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 1 0 0 1 0 0 1 0 1 0 1 0 1 0 1 1 0 0 0 0 0 1 0 0
0 1 0 0 0 0 0 0
B9 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 1 1 0 0 1 1 0 0 1 0 1 0 0 1 1 0 0 1 0 1 0 0
0 1 0 0 1 0 0 0
BA 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 0 0 1 1 0 0 0 0 0 1 0 0 1 0 0 0 0
0 0 0 1 0 0 0 0
BB 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 1 0 0 0 1 1 0 1 0 0 1 1 0 1 1 0 0 1 0 0 1 0 0 1 0 0
0 1 0 0 0 0 0 0
BC 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 1 0 0 1 0 0 1 0 0 0 1 0 0 1 0 1 0 0 0 0 0 0 1 0 0
1 0 0 0 0 0 0 0
BD 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0
0 1 0 0 0 0 0 0

Appendix 6

Similarity matrix from SPH3 data produced using the simple matching coefficient:



Appendix 7

SPH3+7 presence (1) and absence (0) data:

T1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0
1 0 0 0 0 1 0 1 0 0 0 0 0 1 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
T2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 1 0 1 0 1 0 0 0
0 0 0 0 0 1 0 1 0 1 1 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
T3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 1 0 0
0 1 0 1 0 1 0 1 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
T4 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 1 1 0 0 0 0 0 1 0 0 0 0 0 0 1 0 1 0 1 1 0 1 0 1 0 0 0 1 1
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
T5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 1 0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 1 0 0 0 1 0 1 0 0
0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
T6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 1 0 0
0 0 0 1 0 1 0 1 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
T7 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 1 0 1 1 0 0 0 1 0 1 0 1 1 1 0 0 0 0 1 0 0 0 0
0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 1 0 1 0 0 0 1 0 0 0 0 0 0
0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
T8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 0 0 1 1 0 0 0 1 0 0 0 0
0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0
0 0 0 0 1 0 1 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
T9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0
0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
TA 0 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0 0 1 1 1 0 1 0 0 0 0 1 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
TB 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0 1 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 0 0 0 1 0 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0
0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

TC 0000000000000000000000000010000001100010001001000010001
0000000000001000010001001000010000100001010001000100000
0100100100100000000000000000000000000000
TD 0000000000010000000000001001100100010111010010000001010
0000000000000000000000000000000000000000000000000010000000
0100100000101000000100000000000000000000
TE 0000000000000000000000010001001000110010010000000000000
0000000000000000000000000010000100001000000001001001001 0
0000100110010000000101000000000000000000
TF 0000000000010000001000100100001000100010100100001000 0
0000000000000000010000001000001001000100010000100001000
0000010100100000000000000000000000000000
TG 0000000010101001000010000100011011010001001000010000
0000000100000001010000000001000100100000010000010010 00
0000010100010000100000000000000000000000
TH 0000000001000011010100100010100111010001011010010010
0000000000000000001000000000100100100100011000110101 00
0001010101100000000000000000000000000000
TI 0000000000000000000000001001001101010100001010000001010
0000100000000000000000000000001000100010001000100001000
0100010000000000000100100000000000000000
V1 0000000000000000000000001000100000101001010001100101000
0001000000000000000000000000001000100000000100100000010
0001001010001000011000000000000000000000
V2 0000000001000001000010000001001100100000100010001001 0
0000000000000000000000000000000000000000000000000000000000
0000000000001000000000000000001 0000000
V3 0000000000000000100000100010001001100101001000010100 0
0001000000000000000000000000001000100001000111001001011
0000010010001001101000000001000001 00000
V4 0000000000001000000010010010001000010100001000000010
0000000000000000000000000000001000100001000100000010100
0000010100001010001000000000000001000000
V5 0000000000000000000000000001000000000000000000000000000
0000000000000000001000000000100010000000100001000100
0000010100101000000000000000000001000000

V6 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 1 0 0 1 1 0 0 0 1 0 1 0 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 1 0 0 0 0 1 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 1 0 0 1
0 0 0 0 0 0 1 1 0 0 0 1 1 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0
V7 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 0 0 1 0 0 0 1 1 1 0 0 1 0 1 0 0 0 1 1 0 0 0 1 0 1 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 1 0 0
0 0 0 0 1 1 0 1 0 0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
V8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 1 1 1 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 1 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0
0 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0
V9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 1 0 0 0 0 0 1 0 1 0 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0
0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1
VA 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 1 0 0 0 0 0 1 0 1 0 0 0 1 0 1 1 0 0 1 0 0 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 1 0 1 0 0 1 0 0 0 1 0
0 0 1 0 0 0 1 0 0 1 0 0 0 1 0 1 0 0 1 0 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0
VB 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 1 0 0 0 0 1 0 1 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0
0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0
VC 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 1 0 1 0 1 0 1 0 0 0 0 1 0 0 0 1 0 0 1 0 0 1 1 0 1 0 0 0 0 0 1 0 0
0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 1 0
0 0 1 0 0 1 0 1 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0
VD 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 0 0 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 1 0 0 0 0 0 0 1 0 1 0 0 1 0 0 1 0 0 0 1 1 0 0 0 0 1 0 0 1 0 1
0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0
VE 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 0 1 1 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 0 1 1 0 0 0 1 0
0 0 0 0 1 0 0 1 0 1 0 0 1 1 0 0 1 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0
VF 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 1 0 0 0 1 0 1 1 1 0 1 0 0 0 0 1 0 1 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1
0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
VG 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 1 0 0 0 1 0 1 0 1 0 0 0 1 0 1 1 1 0 1 0 0 1 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 1 0
0 0 1 0 1 0 0 1 0 1 0 0 1 0 1 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
VH 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 1 0 1 1 0 0 1 1 1 0 1 0 1 0 0 0 0 1 0 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 1 1 0 1 0 1
0 0 0 1 0 0 0 1 0 0 1 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0

VI000000000001000000000000100000100100101011000100010010
000100000000000010100000000101100001001001001000010010
000010010100000000110000000000000000000
P1000000001000001000010000100001011000010100100010100 0
000000001010000001010000000101000000000100100000100 10
000100100000000000100000000000001000000
P2000000000000000000000010100100011100110101101001010 00
000000000000010000000000000001001001010001000100001000 0
000101100001000000000000000000000001000000
P3000000000000000110000100000100000100011001011000001 01
000000000000000000000010000100010001000100000100001001 00
001000100000000000010000000000100000000
P4000000001001010100000101001001001001111111000000001000
101000000000000000000000000001000000000010010000100000 0
000100000000000000000000000000010000000
P5000000001000001010100000100100101000100101000101000 0
000000000000000000000000000010100010000010010000010010 0
000100100011000000000000000000010000000
P6000000000000010000010010100100100100010010100000100 00
100100001000000001010010100010100010010001010010010 10
000100000001010101100000000000000000000
P7000000000000001000100100010100110101011110100010010 00
100000000000000000000000010000100010000100000000000000
000100100000000000000000000000010000000
P8000000000000001001010010100110100101010010101000000 10
000000000000000000000000000000000000010000100000000000
010100110000000000000000000000010000000
P9000000000000001000110101001100001000110101100101000 0
010000000010000000000000010001010010000101001000010 10
010010101001010010000000000000010000000
PA**********
********0000000000001000000001000100010001000001000000
000000110000000000001000000000010000000
PB000000010001010100001000001001001000101110100101000 0
010000000000000000001000100000100000100000100000101 01
000100101000010010000000000000000000000

PC0000000000000100010001001010011010101001011001010000
0000000000000000000000000000100100010000100010000010001
0001001000010000010000000000000010000000
PD0000000000100100000001000000000011010001010111000000000
0001000000000001001000000000001000000001010000000001001
0000100000010000001000000000000000000000
PE0000000000100100100100000010001101010101010100000000
0001000000100001001101000100001000100000101000100011001
0000001000000100001000000000000010000000
PF0000000000000001001001010101000110101100101000000000
0000000000000000000001000000000100000010010010000010000
0001000000000000000000000000000010000000
PG0000000000000000001000000010100100010010111111100010000
0000000000000000000000000000001010001000010001000010000100001
0000101000001000100000000000000010000000
PH0000000000000100010000100101001101000111111001010011 0
0000000000000000000000100000100000000000000000000000000000
0000000000000000000100000000001000000000
PI0000000000001000001001010000001010000000001000010 1000
0000000000000000000000011000010010000100010100000000001
0001001000010000000000000000000010000000
M10000000000000000000001001001001001000110101000000110
0001000000000000000000000000000000000001000000000000000
0000000001000000000000000000000000000000
M20000000000000011010001100100101000101100001001001001010
0101000000000000000000000000000000000001000000000010000
0000010001000000000000000000000000000000
M30000000100000010000000010101010011000011000000100010
0000000000000000100000000000001000100000100100010000000
0000000001000000000000000000000000000000
M40000000000000000000000000000001010001000000001000000
0000000000000010000000000000000000000000000000000000000
0000000001000000000000000000000000000000
M50000000010010000000000101011011101000101111001000110
0100001000000000000001000010100010000010000000000000
1000000001000000100000000000000000000000

M6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 1 1 0 1 0 0 1 1 0 1 1 1 0 1 0 1 0 0 0 0 1 1 0
1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0
0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
M7 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 1 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 1 0
1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
M8 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0 1 1 0 0 1 0 1 0 0 1 1 0 0 0 0 0 1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0
0 0 0 0 0 1 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
M9 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 0 0 1 0 1 0 0 0 1 0 1 1 0 0 1 0 1 1 0 1 0 0 1 0 0 1 0
1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0
0 0 0 0 0 1 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
MA 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 1 0 1 0 0 1 0 0 1 0 1 0 0 1 1 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0
0 0 0 0 1 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
MB 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 1 0 1 0 1 0 0 1 0 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 1 0 1 0 0 0 0
0 0 0 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
MC 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 1 1 0 1 0 0 0 0 1 1 0
1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 1 1 0 0 0 0 0 0 0 0 1 0 0 0 0
0 0 0 0 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
MD 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 1 0 1 0 0 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1 0 1 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1 0 0 0 1 1 0 0 1
0 0 0 0 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
ME 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0 1 0 1 1 0 0 0 1 0 1 1 0 0 0 1 0 1 0 0 1 1 0 0 0 0 0 0 0 1 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
MF 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 1 1 0 1 0 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 1 0
0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
MG 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0
0 0 0 0 1 0 1 0 1 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
MH 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

B1 0 0 0 0 0 1 0 0 0 1 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 1 1 0 0 0 1 1 0 1 0 0 1 0 0 1 0 1 1 0 0 0 1 0 0 1 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0
B2 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0
0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B3 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0
1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0
0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B4 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 0 1 0 1 1 0
1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0
0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B5 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 1 0 1 0 0 0 0 1 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0
0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0
B6 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0
0 0 0 0 0 1 0 0 1 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0
B7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 1 0 0 0
0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0
B8 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 1 0 0 1 0 0 1 0 1 0 1 0 1 0 1 1 0 0 0 0 0 1 0 0
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 1 0 1 0 1
0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0
B9 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 1 1 0 0 1 1 0 0 1 0 1 0 0 1 1 0 0 1 0 1 0 0
0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0
0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0
BA 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 1 1 0 0 0 0 0 1 0 0 1 0 0 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 1 1 0 0 0 0 0 0 0 0 1 0 0 1
0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0
BB 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 0 0 1 1 0 1 0 0 1 1 0 1 1 0 0 1 0 0 1 0 0 1 0 0
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0
0 1 0 0 0 1 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0
BC 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 1 0 0 1 0 0 1 0 0 0 1 0 0 1 0 1 0 0 0 0 0 0 1 0 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 0 1 0 0 0 1 0 0 0
0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0

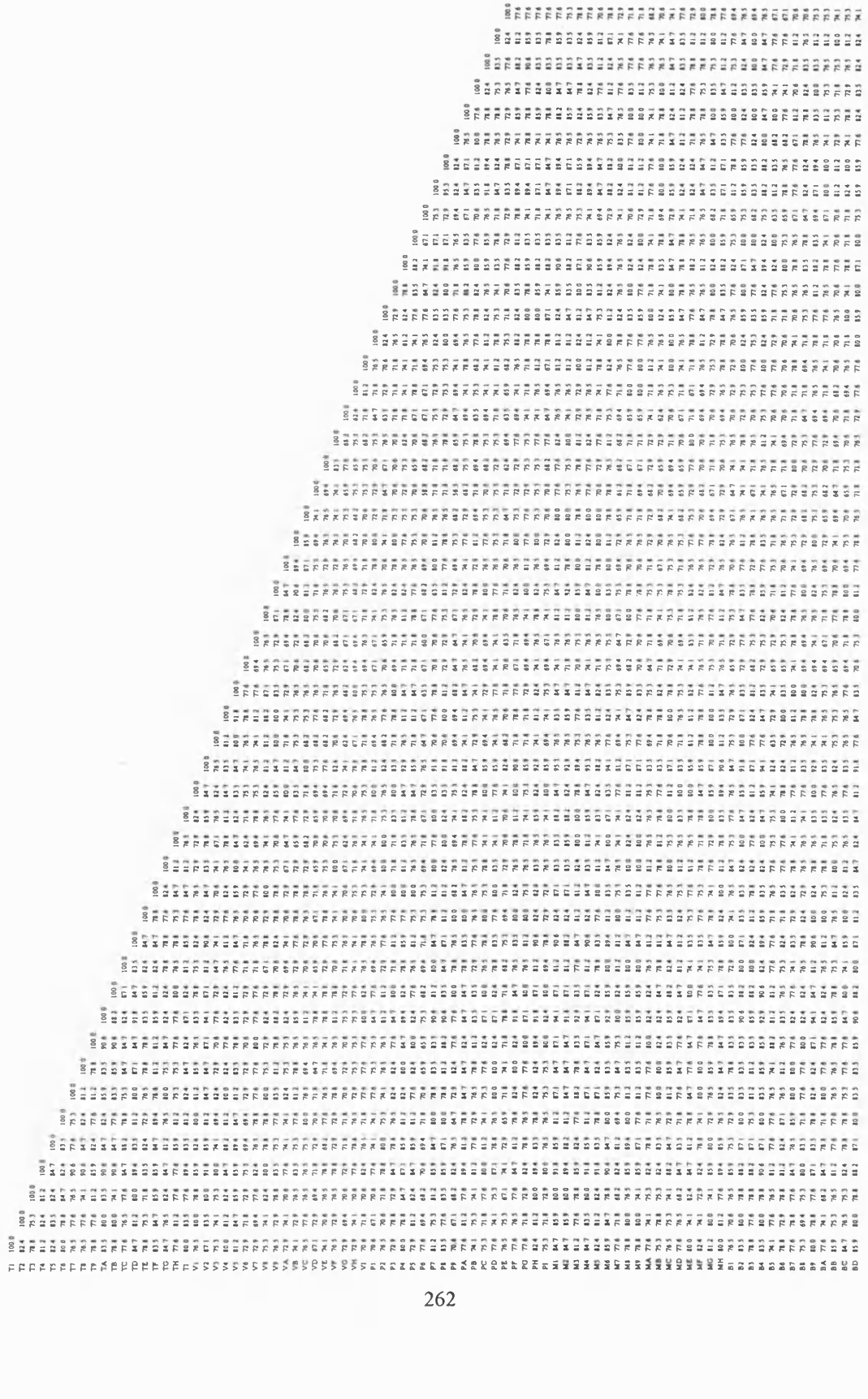
BD 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0

0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0

0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0:

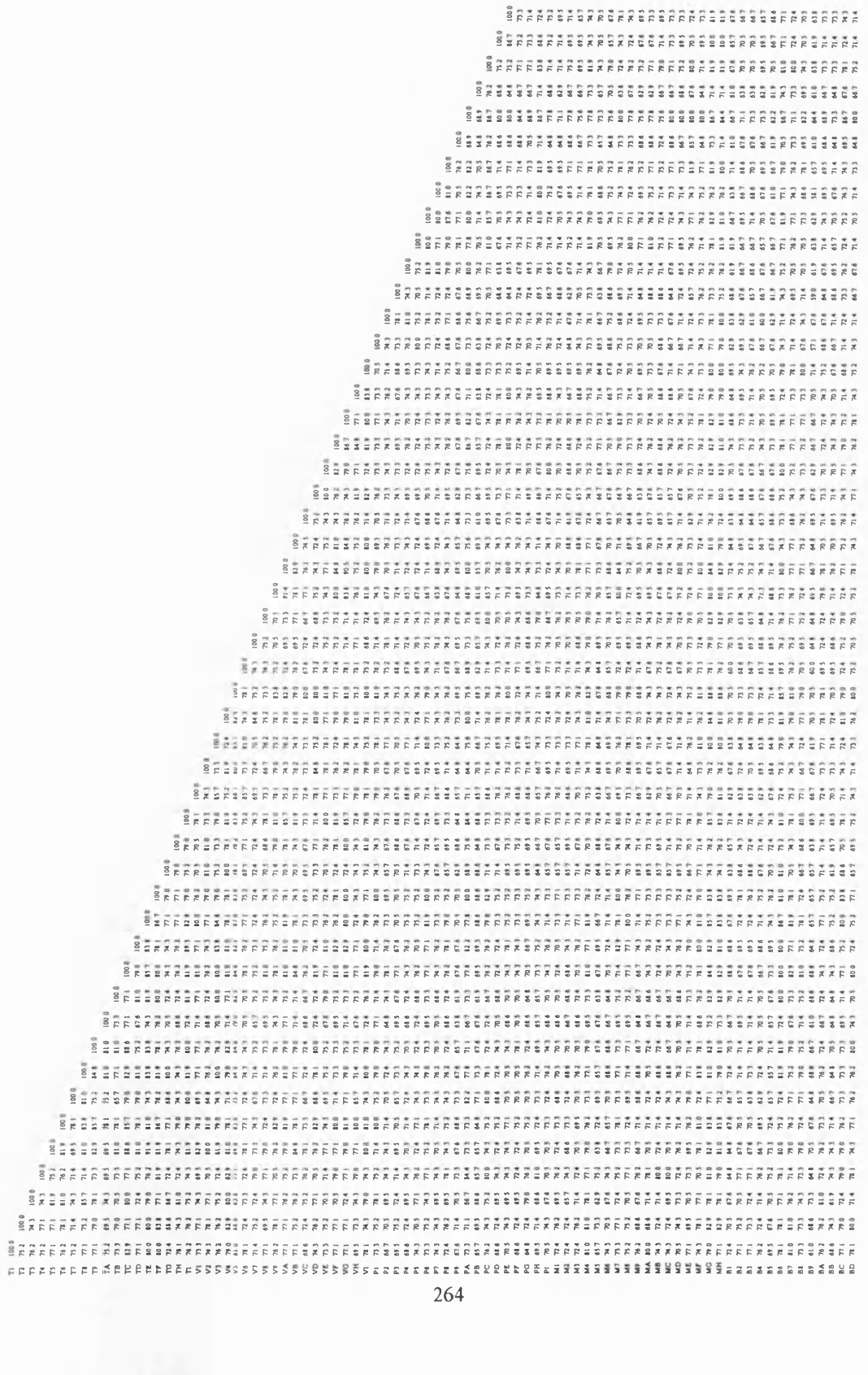
-Appendix 8

Similarity matrix from SPH3+7 data produced using the simple matching coefficient:



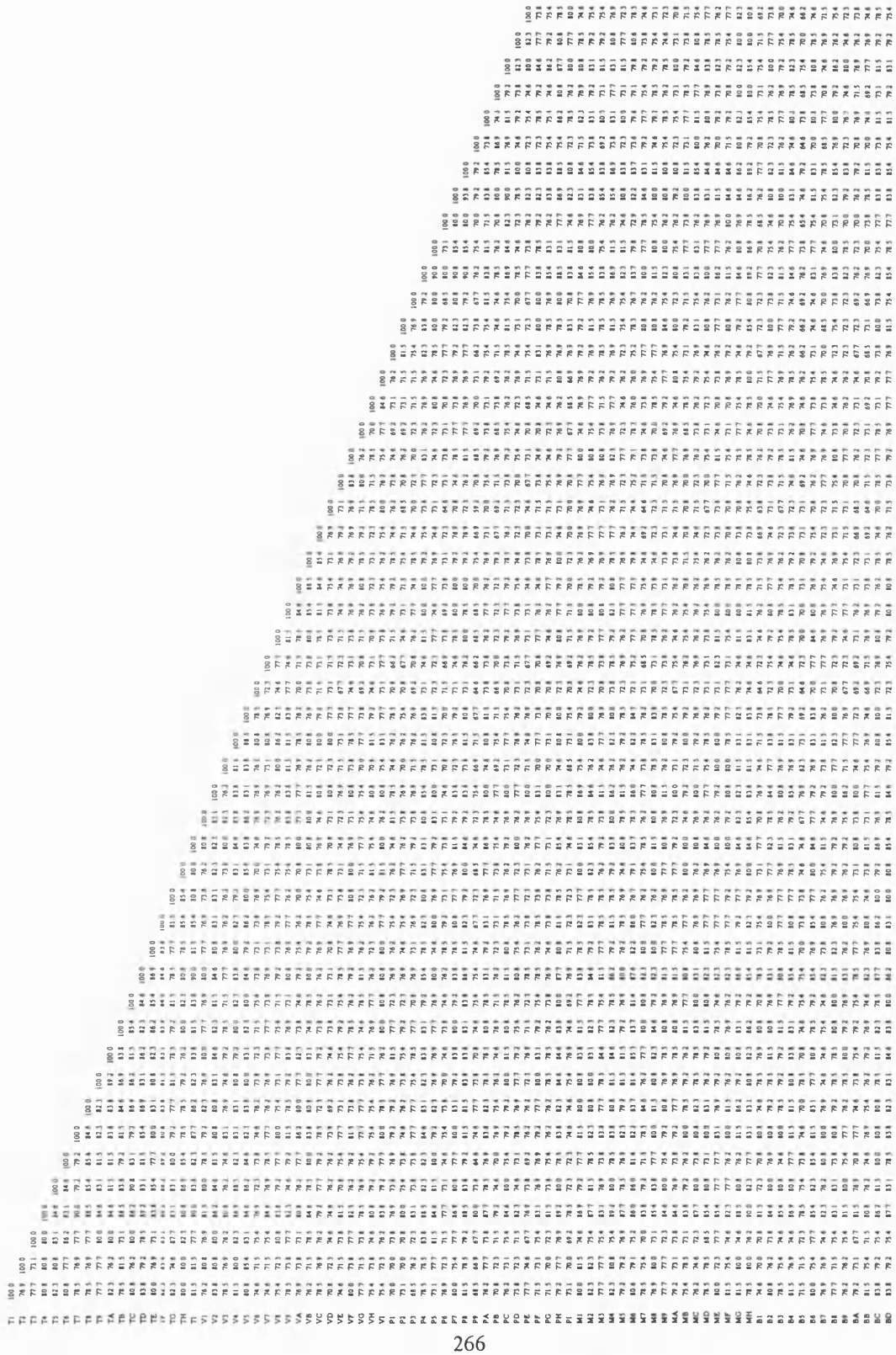
Appendix 9

Similarity matrix from SPH1 & SPH3 data produced using the simple matching coefficient:



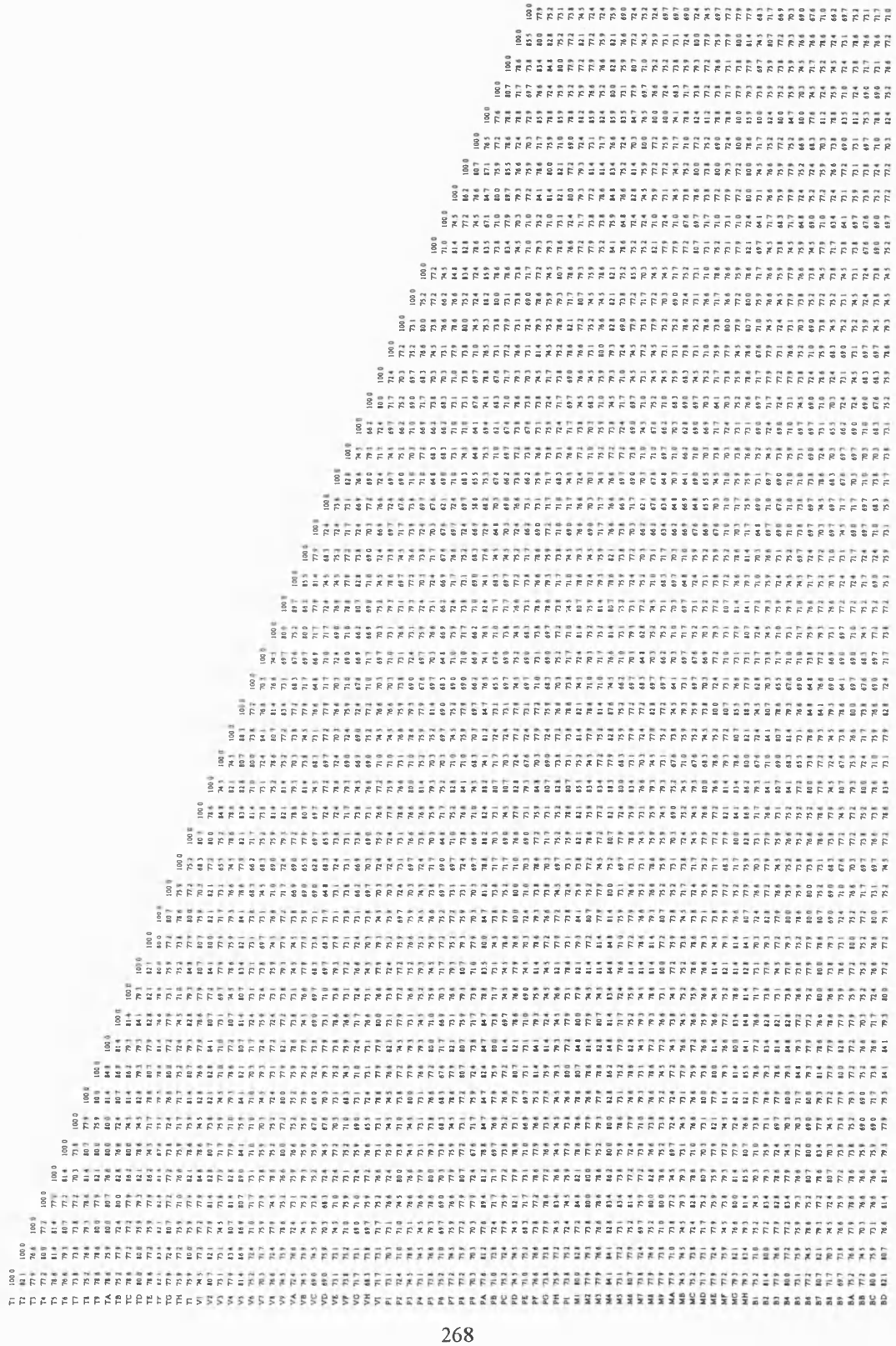
Appendix 10

Similarity matrix from SPH1 & SPH3+7 data produced using the simple matching coefficient:



Appendix 11

Similarity matrix from SPH3 & SPH3+7 data produced using the simple matching coefficient:



Appendix 12

Similarity matrix from SPH1, SPH3 & SPH3+7 data produced using the simple matching coefficient: